INSPECTION TIME IN PATIENTS WITH INTRACRANIAL TUMOURS BEFORE AND AFTER NEUROSURGERY

Jennifer L. Scotland

PhD

The University of Edinburgh

2010

Abstract

Introduction: Many patients with brain tumours experience dysfunction in several cognitive domains. Given the limited survival times of the majority of patients with brain tumours, maintenance or improvement of quality of life is as important as increasing survival time. Impaired cognition has a negative impact on quality of life and as such, cognitive function is becoming an increasingly important endpoint in clinical trials in neuro-oncology. However, measuring cognition in patients with brain tumours is problematic for a number of reasons. Most intracranial tumours are initially treated with surgery and studies of neurosurgical morbidity often evaluate physical as opposed to cognitive domains, yet the latter can have a greater negative impact on the patient's quality of life. This thesis therefore details cognition in brain tumour patients at the time of presentation (pre-operatively) and examines the effects of surgical intervention on cognitive function. Of particular interest is the potential utility of inspection time, a computer-based measure of the brain's information processing efficiency, as a measure of brain slowing as a result of the tumour and as an indicator of response to surgical intervention.

Methods: The study is based on a cohort of 118 newly-presenting patients with a supratentorial brain tumour who were to have surgery (biopsy or resection). Each patient was administered a comprehensive battery of cognitive tests prior to surgery (baseline). The battery comprised inspection time testing, other standardised cognitive measures and assessment of mood, quality of life and functional status. Post-operatively, each patient repeated the inspection time test in addition to a selected number of the other tests administered at baseline. For comparison, a group of patients admitted for elective spinal surgery (n = 85) were also tested pre- and post-operatively. A group of healthy volunteers provided a second control group by being tested twice (n = 80).

Results: The brain tumour cohort were significantly impaired by comparison with both control groups at baseline (pre-operatively) on the majority of the cognitive measures, including inspection time. Baseline inspection time scores were significantly related to some scores on the EORTC Quality of Life Questionnaire in the brain tumour group, but not in the spinal surgery group. There was no significant difference between the brain tumour and spinal surgery groups in term of the levels of pre-operative anxiety and depression.
The brain tumour cohort showed significantly greater relative deterioration on inspection time following surgery by comparison with both control groups. The brain tumour cohort also deteriorated significantly on several other measures post-operatively by comparison with the healthy control group.
Detailed analyses were carried out to determine the differential effects of tumour type, location, and type of surgery (biopsy or resection) on inspection time and other functions in the brain tumour group.

Conclusions: Tumour-related cognitive impairment appears to be common in a heterogeneous group of brain tumour patients with a variety of different tumours located throughout the brain. Surgical intervention has a negative impact on function in brain tumour patients, although this deterioration may be transient. General slowing of visual information processing appears to be common to brain tumour patients and the inspection time task provides a feasible and useful method of assessment in brain tumour patients. The task is sensitive to tumour-related brain slowing and can provide a reliable assessment of response to surgery. Given the task's advantages over more commonly-used cognitive measures, it could be usefully incorporated into cognitive tests batteries in neuro-oncology.

Declaration

I, Jennifer L. Scotland, declare that this thesis is my own work.

I carried out participant recruitment and testing and scored, prepared and analysed data obtained from each testing session. The analyses reported in Chapter 9.5 were carried out by Mike Allerhand, Statistician.

This work has not been submitted for any other degree or professional qualification. Some of the findings detailed in Chapter 4 have been published (Scotland et al., 2009) and the paper is included as an appendix.

Jennifer L. Scotland

7th May 2010

Acknowledgements

Firstly, I would like to thank my supervisors Professor Ian Whittle and Professor Ian Deary for their helpful advice and support. I am particularly grateful to Professor Whittle for his enthusiasm and assistance during the recruitment phase and to Professor Deary for his useful advice and guidance during the analysis phase. I also thank both supervisors for their helpful comments and encouragement during the writing up of this thesis.

I would also like to acknowledge the patients who gave up their time to participate in the study, many of whom were willing to take part during a time of illness and stress. Without their participation, this thesis would not have been possible. I am also grateful to the members of staff on wards 32 and 33 at the Western General Hospital, who provided assistance in identifying potential recruits and particular thanks goes to Shanne McNamara for her enthusiasm and help with patient recruitment. I also thank Roy Welensky for granting permission to use the recruitment panel, and all the volunteers who were willing to take part in the study.

I also wish to extend my gratitude to the Chief Scientist Office for providing funding for the study. I also acknowledge the university IT staff who installed and set up the inspection time programme and would like to thank Mike Allerhand for his assistance with analysing the session 3 inspection time data.

I have been lucky enough to have been surrounded by friendly and supportive colleagues in my daily work and could not have asked for a better group of people to work alongside. I am grateful to each of them for their constant support and belief in me.

I am, as ever, indebted to my parents for their continued encouragement and invaluable support throughout my University career. I also wish to thank my brother for keeping me going with his humour! Finally, I would like to thank my husband, Ally Balfour, for his unconditional love and support and for bearing the brunt of my frustrations. It would not have been able to complete this thesis without him.

Contents

page number

Abstract of thesis ii
Declaration iv
Acknowledgements v
List of Tables xiv
List of Figures xvi
List of Appendices xxi

1 Brain Tumours and Cognition 1

1.1 Brain tumours: an overview 1
1.1.1 Primary brain tumours 1
1.1.2 Secondary brain tumours 3
1.1.3 Risk Factors 4
1.1.4 Common Symptoms 4

1.2 Brain tumours and cognitive function 5

1.3 Tumour type and cognitive function 7
1.3.1 High-grade gliomas 8
1.3.2 Low-grade gliomas 15
1.3.3 High- and Low-Grade Gliomas – Conclusions 20
1.3.4 Other tumour types 20
1.3.5 Tumour Type and Cognition: Conclusions 23

1.4 Brain tumour location and cognition 24

1.5 Treatment effects on cognition 28
1.5.1 Surgery and cognitive function 28
1.5.2 Radiotherapy and cognitive function 32
1.5.3 Prescribed Medication and Cognition 35

1.6 Cognitive function as a prognostic indicator 37

1.7 Issues associated with measurement of cognitive function in brain tumour patients..... 41

1.8 Brain Tumours and Information-Processing 45

2 Inspection Time.....48

2.1 History of Inspection Time 48

2.2 Method of Testing..... 49

2.3 Theory of Inspection Time 51

2.4 Inspection Time and Higher Mental Abilities..... 52

2.4.1 Inspection time and factors of intelligence..... 57

2.5 Neuroanatomical Substrates of Inspection Time 61

2.5.1 Functional MRI Scanning Studies..... 61

2.5.2 Neurochemical Correlates..... 64

2.6 Inspection Time in Children 66

2.7 Inspection Time in Older Adults..... 69

2.8 Inspection Time and Heritability..... 73

2.9 Inspection Time in Clinical Populations 76

2.9.1 Inspection time and Dementia..... 77

2.9.2 Inspection time and Parkinson's disease..... 77

2.9.3 Inspection time and Depression 79

2.9.4 Inspection time and Schizophrenia..... 80

2.9.5 Inspection time and Anaesthesia..... 81

2.9.6 Inspection time and Dyslexia..... 81

2.10 Inspection Time – Conclusions..... 82

2.11 Aims of Thesis 83

3 Methodology 85

3.1 Participants 85

3.2 Patient recruitment 85

3.3 Healthy control group recruitment 86

3.4 Procedure 87

3.5 Measures 89

3.5.1 Cognitive Assessments 89

3.5.2 Mood Assessment 95

3.5.3 Functional Assessments: Edinburgh Functional Impairment Tests (EFIT) 95

3.5.4 Quality of Life Assessment 97

3.5.5 Clinical Assessments 98

4 Recruitment Bias 99

4.1 Overview 99

4.2 Results 99

4.3 Discussion 104

5 Baseline/Pre-Operative Function 108

5.1 Descriptive Data 108

5.1.1 Demographic Information 108

5.1.2 Attrition and Follow-Up 109

5.2 Baseline Analyses: Cognitive Measures 112

5.2.1 Overview of Analysis Procedure 112

5.2.2 Inspection Time Scores: All Inspection Time Data 113

5.2.3 Inspection Time Scores: Valid Inspection Time Data 115
5.2.4 Rey Auditory Verbal Learning Test 117
5.2.5 Trail Making Test Part B 118
5.2.6 Verbal Fluency 118
5.2.7 Digit Symbol Coding 119
5.2.8 Letter-Number Sequencing 120

5.3 Baseline Analyses: Functional Measures 121
5.3.1 Overview of Analysis Procedure 121
5.3.2 Williams Delayed Recall Test (EFIT) 121
5.3.3 Nine Hole Peg Test (Right Hand, EFIT) 122
5.3.4 Nine Hole Peg Test (Left Hand, EFIT) 123
5.3.5 Timed Ten Metre Walk (EFIT) 124

5.4 Baseline Analyses: Mood 125
5.4.1 Anxiety Scores 125
5.4.2 Depression Scores 126
5.4.3 Total Hospital Anxiety and Depression Scale Scores 126

5.5 Baseline Analyses: Functional Measures 127
5.5.1 Overview of Analysis Procedure 127
5.5.2 Barthel Disability Index 127
5.5.3 Karnofsky Performance Scale 128
5.5.4 Boston Aphasia Severity Rating Scale (EFIT) 128

5.6 Discussion 142

6 Post-Operative Function 153

6.1 Overview of Analysis Procedure 153

6.2 Results 154
6.2.1 Inspection Time Scores: All Inspection Time Data 154
6.2.2 Inspection Time Scores: Valid Inspection Time Data 156
6.2.3 Digit Symbol Coding 159

6.2.4 Nine Hole Peg Test (Right Hand, EFIT) 161
6.2.5 Nine Hole Peg Test (Left Hand, EFIT) 163
6.2.6 Williams Delayed Recall Test (EFIT) 165
6.2.7 Timed Ten Metre Walk (EFIT) 168

6.3 Discussion 172

7 The Comparative Effects of Biopsy and Resection 183

7.1 Overview of analysis procedure 183

7.2 Results 184
7.2.1 Inspection Time Scores: All Inspection Time Data 184
7.2.2 Inspection Time Scores: Valid Inspection Time Data 186
7.2.3 Digit Symbol Coding 187
7.2.4 Williams Delayed Recall Test (EFIT) 189
7.2.5 Nine Hole Peg Test (Right Hand, EFIT) 191
7.2.6 Nine Hole Peg Test (Left Hand, EFIT) 193
7.2.7 Timed Ten Metre Walk (EFIT) 195

7.3 Discussion 198

8 Tumour Type, Location and Lateralisation: Baseline/Pre-Operative Function 201

8.1 Overview of Analysis Procedure 201

8.2 Histological Tumour Type 201
8.2.1 Inspection Time Scores: All Inspection Time Data 202
8.2.2 Inspection Time Scores: Valid Inspection Time Data 204
8.2.3 Rey Auditory Verbal Learning Test 205
8.2.4 Trail Making Test Part B 207
8.2.5 Verbal Fluency 208
8.2.6 Digit Symbol Coding 210
8.2.7 Letter-Number Sequencing 211

8.2.8 Williams Delayed Recall Test (EFIT) 213
8.2.9 Nine Hole Peg Test (Right Hand, EFIT) 214
8.2.10 Nine Hole Peg Test (Left Hand, EFIT) 216
8.2.11 Timed Ten Metre Walk (EFIT) 217
8.2.12 Hospital Anxiety and Depression Scale 219

8.3 Tumour Location – Lobe 224
8.3.1 Inspection Time Scores: All Inspection Time Data 224
8.3.2 Inspection Time Scores: Valid Inspection Time Data 226
8.3.3 Rey Auditory Verbal Learning Test 228
8.3.4 Trail Making Test Part B 229
8.3.5 Verbal Fluency 231
8.3.6 Digit Symbol Coding 233
8.3.7 Letter-Number Sequencing 234
8.3.8 Williams Delayed Recall Test (EFIT) 236
8.3.9 Nine Hole Peg Test (Right Hand, EFIT) 238
8.3.10 Nine Hole Peg Test (Left Hand, EFIT) 239
8.3.11 Timed Ten Metre Walk (EFIT) 241
8.3.12 Hospital Anxiety and Depression Scale 242

8.4 Hemispheric Lateralisation 250
8.4.1 Inspection Time Scores: All Inspection Time Data 250
8.4.2 Inspection Time Scores: Valid Inspection Time Data 251
8.4.3 Rey Auditory Verbal Learning Test 251
8.4.4 Trail Making Test Part B 252
8.4.5 Verbal Fluency 253
8.4.6 Digit Symbol Coding 254
8.4.7 Letter-Number Sequencing 254
8.4.8 Williams Delayed Recall Test (EFIT) 255
8.4.9 Nine Hole Peg Test (Right Hand, EFIT) 256
8.4.10 Nine Hole Peg Test (Left Hand, EFIT) 257
8.4.11 Timed Ten Metre Walk (EFIT) 257
8.4.12 Hospital Anxiety and Depression Scale 258

9 Tumour Type, Location and Lateralisation: Post-Operative Function 263

9.1 Overview of analysis procedure 263

9.2 Histological Tumour Type 264

9.2.1 Inspection Time Scores: All Inspection Time Data 264

9.2.2 Inspection Time Scores: Valid Inspection Time Data 266

9.2.3 Digit Symbol Coding 268

9.2.4 Williams Delayed Recall Test (EFIT) 270

9.2.5 Nine Hole Peg Test (Right Hand, EFIT) 272

9.2.6 Nine Hole Peg Test (Left Hand, EFIT) 274

9.2.7 Timed Ten Metre Walk (EFIT) 276

9.3 Tumour Lobe 280

9.3.1 Inspection Time Scores: All Inspection Time Data 280

9.3.2 Inspection Time Scores: Valid Inspection Time Data 282

9.3.3 Digit Symbol Coding 284

9.3.4 Williams Delayed Recall Test (EFIT) 286

9.3.5 Nine Hole Peg Test (Right Hand, EFIT) 288

9.3.6 Nine Hole Peg Test (Left Hand, EFIT) 290

9.3.7 Timed Ten Metre Walk (EFIT) 292

9.4 Hemispheric Lateralisation 296

9.4.1 Inspection Time Scores: All Inspection Time Data 296

9.4.2 Inspection Time Scores: Valid Inspection Time Data 298

9.4.3 Digit Symbol Coding 300

9.4.4 Williams Delayed Recall Test (EFIT) 302

9.4.5 Nine Hole Peg Test (Right Hand, EFIT) 303

9.4.6 Nine Hole Peg Test (Left Hand, EFIT) 305

9.4.7 Timed Ten Metre Walk (EFIT) 306

9.5 Session 3 Follow-Up: Low-Grade Glioma vs. High-Grade Glioma310
9.5.1 Rationale and Overview of Analysis Procedure 310
9.5.2 Results 310

9.6 Tumour Characteristics: Discussion 313
9.6.1 Tumour Histology 313
9.6.2 Hemispheric Location 318
9.6.3 Tumour Location: Lobe 320

10 Quality of Life 324

10.1 Method of Scoring 324

10.2 Quality of Life: Brain tumour patients vs. spinal surgery controls 325

10.3 Brain Cancer Module: Brain tumour patients vs. spinal surgery controls 329

10.4 Quality of Life and Inspection Time 332

10.5 Brain Cancer Module and Inspection Time 332

10.6 Discussion 336
10.6.1 Brain Tumour vs. Spinal Surgery Patients 336
10.6.2 Quality of Life and Inspection Time Scores 338

11 Conclusions and Future Work 341

References 344

List of Tables

Table 3.1. Schedule of tests.

Table 4.1. The number of patients with each tumour type in the total cohort (excluding 2 patients who were recruited but did not undergo surgery to obtain a histological diagnosis), and subgroups of those who participated, declined, were ineligible or missed.

Table 5.1. Demographic characteristics of each study group: brain tumour, spinal surgery and healthy control.

Table 5.2. Crosstabulation showing actual vs. expected counts for qualifications held in each of the three groups (brain tumour, spinal surgery, healthy).

Table 5.3. Histological and location characteristics of the brain tumour group.

Table 5.4. Overview of results of general linear modelling analyses comparing the brain tumour, spinal surgery and healthy control groups on cognitive and mood scales.

Table 5.5. Estimated marginal mean scores, adjusted for age and NART score, for each participant group on each baseline test and post-hoc pairwise comparisons.

Table 6.1. Estimated marginal mean session 2 test scores, adjusted for age, NART score and baseline test score, and post-hoc pairwise comparisons.

Table 8.1. Comparison of histological groups on cognitive and mood scales at baseline.

Table 8.2. Comparison of tumour lobe groups on cognitive and mood scales at baseline.

Table 8.3. Comparison of the left and right hemisphere groups on cognitive and mood scales at baseline.

Table 9.1. Overview of comparisons of post-operative test performance in each histological group.

Table 9.2. Overview of comparisons of post-operative test performance in each tumour lobe group.

Table 9.3. Overview of comparisons of post-operative test performance in the left and right hemisphere groups.

Table 10.1. Mean scores and comparisons of the brain tumour and spinal surgery groups on each EORTC QLQ-C30 scale.

Table 10.2. Mean scores and comparisons of the brain tumour and spinal surgery groups on the QLQ-BN20 scales.

Table 10.3. Non-parametric (spearman's rho) correlations between total baseline inspection time scores and the functional, symptom, and global health scales of the EORTC QLQ-C30.

Table 10.4. Non-parametric (spearman's rho) correlations between total baseline inspection time scores and the symptom scales/items on the QLQ-BN20.

List of Figures

Figure 2.1. The two possible stimulus shapes shown during each inspection time trial.

Figure 2.2. Two possible sequences of events in each inspection time trial.

Figure 3.1. Sequence of events in each inspection time trial.

Figure 4.1. Flowchart detailing brain tumour cohort eligibility for the study and recruitment.

Figure 5.1. Attrition and follow-up rates in the brain tumour, spinal surgery and healthy control groups.

Figure 5.2. Psychometric curves describing inspection time performance (all data) in the brain tumour, spinal surgery and healthy control groups.

Figure 5.3. Psychometric curves describing inspection time performance (valid data only) in the brain tumour, spinal surgery and healthy control groups.

Figure 6.1. Baseline and session 2 inspection time scores for the three participant groups.

Figure 6.2. Baseline and session 2 inspection time scores for the three participant groups including only data from participants with valid scores at baseline.

Figure 6.3. Baseline and session 2 digit symbol coding scores for the three participant groups.

Figure 6.4. Baseline and session 2 nine hole peg test (right hand) scores for the three participant groups.

Figure 6.5. Baseline and session 2 nine hole peg test (left hand) scores for the three participant groups.

Figure 6.6. Baseline and session 2 Williams Delayed Recall Test scores for the three participant groups.

Figure 6.7. Baseline and session 2 Timed Ten Metre Walk scores for the three participant groups.

Figure 7.1. Baseline and session 2 inspection time scores for the biopsy and resection groups.

Figure 7.2. Baseline and session 2 inspection time scores for the biopsy and resection groups with only valid baseline inspection time scores included.

Figure 7.3. Baseline and session 2 digit symbol coding scores for the biopsy and resection groups.

Figure 7.4. Baseline and session 2 Williams Delayed Recall Test scores for the biopsy and resection groups.

Figure 7.5. Baseline and session 2 right hand nine hole peg test scores for the biopsy and resection groups.

Figure 7.6. Baseline and session 2 left hand nine hole peg test scores for the biopsy and resection groups.

Figure 7.7. Baseline and session 2 timed ten metre walk scores for the biopsy and resection groups.

Figure 8.1. Baseline inspection time scores for each histological tumour type group.

Figure 8.2. Baseline inspection time scores including only valid data for each histological tumour type group.

Figure 8.3. Baseline Rey Auditory Verbal Learning Test scores for each histological tumour type group.

Figure 8.4. Baseline trail making test part B scores for each histological tumour type group.

Figure 8.5. Baseline verbal fluency test scores for each histological tumour type group.

Figure 8.6. Baseline digit symbol coding test scores for each histological tumour type group.

Figure 8.7. Baseline letter number sequencing test scores for each histological tumour type group.

Figure 8.8. Baseline Williams Delayed Recall Test scores for each histological tumour type group.

Figure 8.9. Baseline right hand nine hole peg test scores for each histological tumour type group.

Figure 8.10. Baseline left hand nine hole peg test scores for each histological tumour type group.

Figure 8.11. Baseline timed ten metre walk test score for each histological tumour type group.

Figure 8.12. Baseline anxiety scores on the Hospital Anxiety and Depression Scale for each histological tumour type group.

Figure 8.13. Baseline depression scores on the Hospital Anxiety and Depression Scale for each histological tumour type group.

Figure 8.14. Baseline total scores on the Hospital Anxiety and Depression Scale for each histological tumour type group.

Figure 8.15. Baseline inspection time scores for each tumour lobe group.

Figure 8.16. Baseline inspection time scores including only valid scores for each tumour lobe group.

Figure 8.17. Baseline Rey Auditory Verbal Learning Test scores for each tumour lobe group.

Figure 8.18. Baseline trail making test part B scores for each tumour lobe group.

Figure 8.19. Baseline verbal fluency test scores for each tumour lobe group.

Figure 8.20. Baseline digit symbol coding test scores for each tumour lobe group.

Figure 8.21. Baseline letter-number sequencing test scores for each tumour lobe group.

Figure 8.22 Baseline Williams Delayed Recall Test scores for each tumour lobe group.

Figure 8.23. Baseline right hand nine hole peg test scores for each tumour lobe group.

Figure 8.24. Baseline left hand nine hole peg test scores for each tumour lobe group.

Figure 8.25. Baseline timed ten metre walk test scores for each tumour lobe group.

Figure 8.26. Baseline anxiety scores on the Hospital Anxiety and Depression Scale for each tumour lobe group.

Figure 8.27. Baseline depression scores on the Hospital Anxiety and Depression Scale for each tumour lobe group.

Figure 8.28. Baseline total scores on the Hospital Anxiety and Depression Scale for each tumour lobe group.

Figure 9.1. Baseline and session 2 inspection time scores for each histological group.

Figure 9.2. Baseline and session 2 inspection time scores with only valid baseline scores included, for each tumour type group.

Figure 9.3. Baseline and session 2 digit symbol coding scores for each tumour type group.

Figure 9.4. Baseline and session 2 Williams Delayed Recall Test scores for each tumour type group.

Figure 9.5. Baseline and session 2 right hand nine hole peg test scores for each tumour type group.

Figure 9.6. Baseline and session 2 left hand nine hole peg test scores for each tumour type group.

Figure 9.7. Baseline and session 2 timed ten metre walk test scores for each tumour type group.

Figure 9.8. Baseline and session 2 inspection time scores (all data) for each tumour lobe group.

Figure 9.9. Baseline and session 2 inspection time scores (valid baseline data) for each tumour lobe group.

Figure 9.10. Baseline and session 2 digit symbol coding scores for each tumour lobe group.

Figure 9.11. Baseline and session 2 Williams Delayed Recall Test scores for each tumour lobe group.

Figure 9.12. Baseline and session 2 nine hole peg test (right hand) scores for each tumour lobe group.

Figure 9.13. Baseline and session 2 nine hole peg test (left hand) scores for each tumour lobe group.

Figure 9.14. Baseline and session 2 time ten metre walk test scores for each tumour lobe group.

Figure 9.15. Baseline and session 2 inspection time scores for each tumour hemisphere group.

Figure 9.16. Baseline and session 2 inspection time scores (valid baseline data) for each tumour hemisphere group.

Figure 9.17. Baseline and session 2 digit symbol coding scores for each tumour hemisphere group.

Figure 9.18. Baseline and session 2 Williams delayed recall test scores for each tumour hemisphere group.

Figure 9.19. Baseline and session 2 right hand nine hole peg test scores for each tumour hemisphere group.

Figure 9.20. Baseline and session 2 left hand nine hole peg test scores for each tumour hemisphere group.

Figure 9.21. Baseline and session 2 timed ten metre walk scores for each tumour hemisphere group.

Figure 9.22. Baseline, session 2 and session 3 inspection time scores for the low-grade and high-grade glioma groups.

List of Appendices

Appendix A: Recruitment Poster

Appendix B: Patient Information Sheet

Appendix C: Healthy Volunteer Information Sheet

Appendix D: Letter to Recruitment Panel

Appendix E: Patient Consent Form

Appendix F: Healthy Volunteer Consent Form

Appendix G: Order of Test Administration

Appendix H: Rey Auditory Verbal Learning Test Word Lists

Appendix I: National Adult Reading Test Word List

Appendix J: Trail Making Test Part B Test Sheet

Appendix K: Verbal Fluency Recording Sheet

Appendix L: Digit Symbol Coding Test Sheet

Appendix M: Letter-Number Sequencing Test Items

Appendix N: Hospital Anxiety and Depression Scale

Appendix O: Williams Delayed Recall Test A, B and C

Appendix P: Boston Aphasia Severity Rating Scale 'Cookie Theft' Picture

Appendix Q: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire version 3 and Brain Cancer Module

Appendix R: Barthel Disability Index

Appendix S: Karnofsky Performance Scale

Appendix T: Published Paper: Recruitment difficulties in brain tumour patients cause participation bias: findings from a neuropsychological study of adult inpatients with supratentorial intracranial tumours

Appendix U: Chapter – Returners vs. Non-Returners: Comparison of Baseline Scores
U.1. Overview of Analysis Procedure
U.2. Demographic Comparisons
U.3. Baseline Test Comparisons
U.4. Discussion

1 Brain Tumours and Cognition

1.1 Brain tumours: an overview

1.1.1 Primary brain tumours

Primary brain tumours (i.e. tumours that originate in the cells of the brain parenchyma) are fairly uncommon among adult cancer patients, with an approximate incidence of 8/100,000 per year, (Grant, 2004), accounting for around 2% of all cancers in adults (McKinney, 2004). In the United Kingdom, around 4400 people are newly diagnosed with a primary brain tumour each year and, given that survival is generally very poor and many primary tumour types are incurable, the diagnosis of a brain tumour inevitably has a devastating impact on every aspect of the patients' life. The poor survival times generally associated with a diagnosis of a primary brain tumour means that a disproportionate number of years of life are lost to this disease when compared with many other more common cancers (McKinney, 2004).

1.1.1.1 Gliomas

Gliomas arise from the glial cells that form the connective tissue in the brain and comprise > 90% of all primary brain tumours. The incidence of gliomas is around five to ten per 100,000 general population (Behin et al., 2003) and around half of all brain tumours in adults can be classified as gliomas (Lezak et al., 2004). Cerebral gliomas as a group have a 5-year survival rate of only 18% (Grant, 2004). There are four main types of glioma: astrocytomas, oligodendrogliomas, ependymomas and mixed gliomas. These tumours are graded by the World Health Organisation (WHO) according to the most malignant area identified in them and range from highly malignant (grade 3 or 4) to the less aggressive grades 1 or 2 (Fuller, 2008).

Malignant glial tumours are most commonly classified as anaplastic astrocytomas, anaplastic oligodendrogliomas or anaplastic oligoastrocytomas (WHO Grade 3

tumours) or glioblastoma multiforme (GBM; WHO Grade 4 tumours). Anaplastic astrocytomas and glioblastoma multiforme are unfortunately the most common glial tumours in adults (DeAngelis, 2001). Treatment for these aggressive tumours tends to be palliative and comprises surgical resection of as much tumour as is safely possible, followed by focused cranial irradiation. Concomitant chemotherapy may modestly prolong survival in some patients (Stupp et al., 2005) and Gliadel wafers can be placed at the tumour site during surgery to provide local chemotherapy in either newly diagnosed or recurrent malignant gliomas (Westphal et al., 2006). Despite these advances in treatments, median survival for GBM patients in Randomized Controlled Trials is only around 14 months, and for anaplastic astrocytoma patients is around 2 to 4 years (DeAngelis, 2001, Stupp et al., 2005). However, survival times are variable and depend upon a number of different factors including patient age and symptoms at diagnosis as well as tumour location and the extent of resection achieved during surgery. The mean age at onset is 53 years for GBM and 40 years for anaplastic astrocytoma (Behin et al., 2003).

Low-grade gliomas (LGG) comprise approximately 40% of all adult gliomas (Correa et al., 2007) and can be divided into three main subtypes – diffuse astrocytomas, oligodendrogliomas and mixed oligoastrocytomas. Low-grade gliomas are more common in the third and fourth decades of life and the mean age of occurrence is 35 years for astrocytomas and 45 years for oligoastrocytomas (Behin et al., 2003). Like high-grade gliomas, these tumours infiltrate normal surrounding brain tissue. However, low-grade gliomas are more slowly growing than their high-grade counterparts. The best way to manage patients with low-grade tumours has attracted a great deal of debate, particularly given the fact that many patients are relatively symptom-free. Surgery may be carried out to confirm the radiological diagnosis and radiotherapy may be offered immediately, or deferred until such a time as the patient becomes symptomatic (Whittle, 2002). Almost all patients who present with a low-grade glioma will experience tumour progression during the course of their disease, at which point the recurrent tumour will often have transformed into a more malignant subtype. Median survival time for patients diagnosed with a low-grade glioma is approximately 5 years (DeAngelis, 2001). However, survival times for

patients with low-grade gliomas vary considerably, with many younger patients surviving for much longer (Pignatti et al., 2002).

1.1.1.2 Meningiomas

Meningiomas arise from the cells forming the external membranes that cover the brain (the meninges). They are the most common benign intracranial tumour and the second most common primary brain tumour in adults after gliomas, comprising 13-26% of all primary intracranial tumours (Whittle et al., 2004). This tumour type is more common in older people and in women than in men (ratio 2:1). Ninety percent of meningiomas are classified as WHO grade I and are benign. These tumours can often be cured with complete surgical removal since they tend not to invade the brain itself. However, depending on the extent of excision and WHO tumour grade, up to 20% may recur within a 10 year period at which point further surgery and/or radiotherapy can frequently be offered to the patient (Whittle et al., 2004).

1.1.2 Secondary brain tumours

Secondary, or metastatic brain tumours originate from solid malignant tumours located in other organs in the body that have been transported into the central nervous system. The most common cause of cerebral metastases are tumours in the lung, followed by the breast, melanoma, gastrointestinal tract and kidney cancers (Patchell, 2003). Brain metastases are multiple in more than 50% of cases and are often treated with surgery followed by whole-brain radiotherapy, and occasionally chemotherapy if the primary tumour is sensitive to chemotherapy (Lezak et al., 2004, van den Bent, 2003). Even with this aggressive treatment, median survival time for patients with brain metastases is generally only between 4 and 6 months. However this does depend greatly upon the type and status of the systemic cancer, patient age and performance status (Nussbaum et al., 1996).

1.1.3 Risk Factors

The only confirmed risk factor for developing a brain tumour is prior exposure to ionising radiation (McKinney, 2004). A number of other environmental factors have been investigated as possible risk factors for brain tumours and these have included mobile phones, allergies, diet, tobacco and alcohol, chemical agents and prior head injury. However, there is no consistent evidence to suggest a link between any of these factors and subsequent development of a brain tumour (McKinney, 2004).

1.1.4 Common Symptoms

Behin et al. (2003) classify the most common presenting symptoms that lead to diagnosis of brain tumour into four main categories. The first of these is *seizure*, which may be partial or generalised. Seizures are more commonly a symptom of low grade, as opposed to high-grade glioma. The second category comprises progressive headache, nausea, vomiting and/or visual abnormalities that result from *increased intracranial pressure*. Thirdly, *focal neurological deficits*, dependent upon the tumour site, can reveal an underlying brain tumour. These deficits can include impaired motor function, speech problems and/or visual field deficits. Lastly, *cognitive impairment* is often a sign of underlying tumour and is commonly observed in patients with tumours located in the frontal lobes (Behin et al., 2003).

Grant (2004) completed a Scottish audit of the symptoms recorded in 324 patients who had an imaging diagnosis of a single intracerebral lesion that was deemed most likely to be a glioma. Headache was the most common first symptom in this group of patients (23.5%) and, further, was present in 46.5% of patients at the time of admission to hospital. However, despite the high proportion of patients who experience headaches at some point, it is rare for headache to be the sole presenting symptom of a brain tumour. In fact, this was the case for only 10 of the 324 patients studied. Seizures were found to be the next most common presenting symptom, and were recorded in 21% of the patients studied. Focal symptoms are also common in patients with brain tumours and may include dysphasia, hemiparesis or diplopia. These focal signs are less likely to be the presenting symptom; however,

approximately 80% of patients exhibited focal signs by the time of hospital admission and diagnosis. Furthermore, non-focal symptoms such as confusion and personality change are also commonly observed by the patient or their relatives but again, it is rare for this type of symptom to be the presenting one (Grant, 2004).

1.2 Brain tumours and cognitive function

Cognitive function covers the different processes by which sensory input is elaborated, transformed, stored, recovered and used. Lezak et al. (2004) liken the four major classes of cognitive functions to the computer processes of *input*, *storage*, *processing* and *output*. Therefore, *receptive functions*, are the 'input' functions that include the ability to select, acquire, classify and integrate information. *Memory and learning* are the 'storage' processes by which information is stored and retrieved. *Thinking* is the 'processing' function by which information is organised and *expressive functions* are the 'output' processes by which information is expressed, communicated and acted upon (Lezak et al., 2004). Thus, cognitive functions are those brain functions that are extremely highly-developed in human beings as compared to other mammals (Taphoorn and Klein, 2004). Gilroy (2000) divides cognitive function into nine primary domains and these include attention, concentration, visuospatial skills, language, memory, executive function and sensory and perceptive abilities.

In addition to the aforementioned common symptoms described by Grant (2004), problems with cognitive functioning are frequently observed in patients with brain tumours, and have in fact been cited by Boake and Meyers (1993) as the most common neurologic problem associated with this disease. However, the diagnosis of a brain tumour is rarely made as a result of reduced cognitive functioning alone and, more commonly, impairment in cognitive ability is recognised by the patient and/or relatives retrospectively (Tucha et al., 2000).

There exist a number of reasons for which studying the neurocognitive function of patients with brain tumours is important. Meyers and Brown (2006) highlight the fact

that studying cognition in brain tumour patients enables us to determine exactly what cognitive problems exist prior to any treatment. This not only facilitates the use of effective interventions and guides treatment decisions, but also establishes a baseline level of function by which the effects of treatment can be examined.

Cognitive impairment in one or many aspects of function can have a significantly negative effect on an individual's quality of life, preventing them from taking part in work, leisure and even activities of daily living. Reduced cognitive function not only affects the patient but can have a significantly negative impact on the patient's carer. Therefore, in addition to measures of progression-free survival and objective response on CT or MRI scanning, cognitive functioning is becoming an increasingly important outcome measure in clinical trials in neuro-oncology since it can provide meaningful information with regard to the clinical status of the patient. Furthermore, measuring cognition in brain tumour patients can yield important information regarding side-effects and potential toxicity of selected treatments. Given the limited survival times for many brain tumour patients, cognitive assessment may also reveal a beneficial treatment to be one that stabilises or even slows the progression of worsening mental symptoms, regardless of whether or not overall survival is extended (Meyers and Hess, 2003). Thus, in-depth assessments of cognitive function are now often used in clinical trials, as well as the more commonly used measures of performance status, such as the Karnofsky Performance Scale (KPS). The KPS classifies patients according to their level of functional impairment and allows medical staff to score a patient in terms of their ability to carry out activities of daily living. However, in contrast with detailed cognitive assessment, the KPS, alongside many other measures of performance status, have been shown to have poor validity and reliability and do not relate significantly to measures of quality of life (Meyers and Brown, 2006).

Significant proportions of brain tumour patients have been found to experience cognitive impairment and the reported prevalence of cognitive disorder ranges from around 29% in low-grade glioma patients who have not received radiotherapy to as large a proportion as 90% in patients with diverse brain tumours (Tucha et al., 2000,

Klein et al., 2001). The variation in the reported prevalence of cognitive impairment is the result of a number of factors including different groups of patients studied, different cognitive tests used, treatment factors and different normative data used to compare test scores (Gehring et al., 2008). Studies have been unable to elucidate exactly what causes the observed cognitive impairment in brain tumour patients but it is likely that these deficits arise as a result of a number of different factors. For example, cognitive function in brain tumour patients may be negatively affected by the tumour itself, by tumour-related epilepsy, by treatment factors, and/or by patient-related factors such as age and anxiety or distress (Bosma et al., 2007). This chapter will discuss the studies that have examined cognition in brain tumour patients and will specifically consider: the effects of different tumour types on cognition; the effects of treatments such as medications, neurosurgery and radiotherapy; and will examine evidence to suggest that tumours in a particular location may result in specific cognitive deficits. The role of cognitive assessment in providing an indication of tumour progression will also be discussed alongside the specific problems that are inherently associated with measuring cognition in brain tumour patients.

1.3 Tumour type and cognitive function

As highlighted above, brain tumours may be divided into different groups dependent on the cells from which they arise, how invasive they are and how aggressive they tend to be. Given that brain tumours are not a single disease and that different types of tumour affect the brain and surrounding tissues in different ways, many studies have focused on particular groups of brain tumour patients in order to determine the specific effects of different tumour types on cognitive function (Ek et al., 2005). Determining the specific contribution of the brain tumour itself to cognitive impairment is problematic since many studies include patients whose scores on cognitive measures are confounded by several variables, including the potential effects of radiotherapy and/or chemotherapy. However, some studies have attempted to examine the specific effect of the tumour itself on cognition, by assessing patients prior to any surgical or other intervention and comparing test scores with control

participants or with population norms. Much of this research has focused on the specific cognitive effects of high and low-grade gliomas. However, a smaller number of studies have addressed the cognitive effects of brain metastases and meningiomas. Each tumour subtype and the evidence regarding its effect on cognition will be discussed in the following sections.

1.3.1 High-grade gliomas

The cognitive effects of these aggressive, malignant high-grade tumours have received the greatest attention and this may simply be because they are the most common tumour and thus patients are more readily available.

Salander et al. (1995), given the scarcity of literature in this area at the time, addressed the extent of possible cognitive impairment in patients with high-grade gliomas. Thirty consecutive patients were recruited into the study along with their partners who served as control participants. Patients completed a short, clinical cognitive assessment test (Standardised Mini Mental State Examination, MMSE) at the time of discharge from the neurosurgical department and also 2 months post-radiotherapy. The MMSE is a brief standardised tool used to measure cognitive function and is frequently used to assess patients with dementia (Brown et al., 2004). Five months after the initial assessment and following radiotherapy, those patients who were still alive and in good enough health, with no signs of tumour recurrence, were tested on a more comprehensive battery of neurocognitive tests. The control group of partners also completed the test battery. The test battery was designed to emphasise measurement of memory function and included the Rivermead test, selective reminding and selected subtests from the Revised version of the Wechsler Adult Intelligence Scales (Digit Span, Digit Symbol, Arithmetic, Vocabulary and Picture Arrangement). Eleven patients from the original cohort of thirty completed the test battery and there emerged a clear trend. No significant differences were found on digit span, action memory or prospective memory. However, patients with high-grade gliomas located in either the frontal or temporal lobe exhibited impairment in long-term memory, as compared to short-term memory with long-term

recall, delayed recall and consistent long-term retrieval significantly poorer in the high-grade glioma group compared to the control group. This was despite the fact that all but one of the patients were considered "normal" at baseline, according to their MMSE scores and this therefore suggests that the MMSE may be insensitive to cognitive impairments in brain tumour patients. This is perhaps unsurprising given that many authors have found the measure to have a large ceiling effect (Hogervorst et al., 2002). However, since these patients completed the test battery after receiving radiotherapy to the brain, this makes it difficult to determine the relative contribution of the tumour itself to the observed cognitive impairment since radiotherapy effects may also have had some effect. The specific effects of radiotherapy on cognitive function will be addressed later in this chapter (see section 1.5.2). There were no significant differences between the high-grade glioma patients and the control group in terms of performance on tests that provide a measure of more global cognitive or intellectual abilities (Digit Span, Digit Symbol, Arithmetic, Vocabulary, and Ravens Progressive Matrices). This suggests that long-term memory is significantly affected by the presence of a high-grade glioma and/or radiotherapy to treat the tumour, yet many other cognitive functions remain intact. However, there exist several methodological difficulties with this study that may question this conclusion. Only eleven patients completed the cognitive assessments on both occasions and, therefore, whether these results may be generalised to the whole population of patients with high-grade gliomas is questionable. It is likely that the patients who completed both testing sessions were a highly select group, comprising the best functioning patients. The test battery reportedly lasted between 2 and 3 hours and this is likely to have been somewhat demanding for patients and as a result some test scores may have been artificially poor as a result of fatigue, poor motivation or stress. Although the authors accept that radiotherapy may have been the main cause of their findings, they claim that the observed memory impairment may also have resulted from the high-grade tumour itself. Had the authors tested patients before radiotherapy treatment had commenced, or compared the study group with a group of patients with low-grade tumours who had received early radiotherapy this would have allowed a better analysis of the relative effect of the high-grade glioma itself in contributing to cognitive dysfunction.

A number of other studies have also assessed the neurocognitive function of patients with high-grade brain tumours. Hahn et al. (2003) prospectively examined neuropsychological function and quality of life in adult patients with newly diagnosed primary high-grade gliomas. An initial testing session was carried out prior to, or shortly after, radiotherapy had commenced in order to eliminate radiotherapy treatment effects and thus assess the impact of the tumour itself on cognition more effectively. A total of 68 patients, each of whom completed a battery of tests measuring different aspects of cognition, were recruited. The tests administered to each participant included the American Nelson Adult Reading Test as a measure of premorbid intelligence, Controlled Oral Word Association as a measure of verbal fluency, the Short-Memory Questionnaire, a 14 item self-report measure that records memory deficits on a 4 point scale, Trail Making Tests A and B as an assessment of speed, attention, sequencing, mental flexibility and motor function, and the Visual Reproduction subtest of the Wechsler Memory Scale (Revised), among others. Thus, the test battery comprised assessments of several different aspects of cognition. In addition to several other analyses which will be discussed in detail later in this chapter, the authors examined the specific effect of tumour histopathology on cognitive function. Analyses, that accounted for the known effects of age on cognitive test performance, revealed statistically significant differences in the neuropsychological test performance of GBM patients (n = 30) and those with other gliomas (n = 38). However, further analysis revealed no significant univariate effects for each cognitive measure with the exception of Trail Making Test A, on which GBM patients performed significantly slower. The number of patients with low-grade glioma (n = 18; 26%) included in the study was relatively small compared with the number of high-grade glioma patients (n = 50; 74%) and the patients in the “other” glioma group with which the GBM patients were compared included both high and low-grade glioma subtypes. Therefore, the impairment in cognitive functioning between high-grade glioma patients and LGG patients as a whole may be more extensive than this study suggests, given that GBM patients were considered separately from other high-grade glioma tumour types.

In a further study in this area, Kayl and Meyers (2003) compared the cognitive performance of patients with GBM and anaplastic astrocytoma. The post-surgical neuropsychological functioning of 24 patients with newly diagnosed GBM and 24 patients with newly diagnosed anaplastic astrocytoma were analysed retrospectively. The two groups of patients were matched in terms of age, gender, education, tumour location and tumour volume and each patient had completed a detailed battery of cognitive tests following surgery but prior to commencing radiotherapy or chemotherapy. The test battery comprised a number of standardised tests assessing several aspects of cognition as categorised by the authors themselves. The battery included assessment of intellectual functioning (7 WAIS –R subtests: Information, Digit Span, Arithmetic, Similarities, Comprehension, Block Design and Digit Symbol), memory (Buschke Selective reminding, Hopkins Verbal Learning Test, Benton Visual Retention Test), language and verbal fluency (Visual Naming, Controlled Oral Word Association, Token Test), motor function (Right and Left Grip Strength, Right and Left Grooved Pegboard) and executive functioning (Booklet Category Test, Trail Making Test Part A and B). No significant effect of tumour type (anaplastic astrocytoma vs. GBM) on test scores assessing any of these five cognitive domains was found. Further analysis suggested that patient age has a greater effect on test performance than tumour histology or tumour volume and the authors concluded that, according to this small study, those with more malignant disease do not suffer greater cognitive impairment than patients with less aggressive tumours. However, again, this study did not include any comparison of performance of patients with low-grade tumours, and the authors themselves highlight the fact that this may have influenced results to some degree. Additionally, given the relatively narrow age range of the participants (25 – 52 years) and the fact that the authors did not make any attempt to measure or control for the influence of differing levels of anxiety on test performance suggests that the study had low power to detect any differences between the groups. Moreover, since patients were recruited in the post-operative period, this study is also unable to distinguish surgery effects from tumour effects on cognition, although it does eliminate the potential effects of radiotherapy.

Klein et al. (2001) also assessed the cognitive functioning of high-grade glioma patients in the post-surgical period before radiotherapy and/or chemotherapy treatments had been commenced. They aimed to overcome limitations associated with previous studies that demonstrated the presence of cognitive deficits in brain tumour patients before the start of radiotherapy (i.e. deficits that are more likely caused by the presence of the tumour itself) but did not recruit adequate control groups (Scheibel et al., 1996, Taylor et al., 1998). Klein et al. (2001) argue that studies examining cognition in brain tumour patients should be based upon comparisons with a control group of patients with a diagnosis of cancer that does not involve the central nervous system in order to control for the many unique stressors encountered by cancer patients. Given that cancer patients are likely to experience significantly increased feelings of anxiety and/or depression, this emotional disturbance could negatively impact upon their performance and lead to inflated estimates of the extent of cognitive impairment in this patient group. This study therefore assessed the neuropsychological functioning and the health-related quality of life of newly diagnosed high-grade glioma patients who had undergone surgery to obtain a histological diagnosis. The high-grade glioma patients (n = 68) were compared with 50 non-small cell lung cancer (NSCLC) patients on a battery of standardised tests designed to assess a wide range of cognitive functions including perception (Line Bisection Test, Facial Recognition Test, Judgement of Line Orientation and Letter-Digit Substitution Test), memory (Visual Verbal Learning Test, Working Memory Task) and attention and executive function (Stroop Colour Word Test, Categoric Word Fluency and Concept Shifting Task). General cognitive performance was also assessed by the Mini Mental State Examination (MMSE) and premorbid intelligence was measured using the Dutch Adult Reading Test. An impairment score was calculated for each individual, since the authors argue that considering only group means may obscure impairments at the individual level. Thus, test scores were converted to z-scores using the mean scores of the NSCLC patients as a reference in order to control for the impact of a cancer diagnosis. Subsequently, a mean overall composite z-score was calculated and impairment was defined as a score of two standard deviations below the mean score of the NSCLC group. For each individual high-grade glioma patient, the number of tests that met

this criterion was counted to obtain an overall impairment score. Since the authors also expected cognitive deficits to be present in the NSCLC group, the fifth percentile of this group was used as a cut-off for neuropsychological impairment. Both glioma and NSCLC patients were deemed cognitively impaired if they had scores that met these criteria on at least three of the cognitive tests administered. Based upon this method of analyses, there was found to be a general reduction in cognitive performance of the brain tumour patients when compared to the NSCLC patients, as follows. Firstly, the HGG patients had significantly lower scores on the MMSE. Moreover, 49% of the glioma group were classified as cognitively impaired, on the basis of the aforementioned criteria, compared with 26% of the NSCLC group and this difference was statistically significant. Furthermore, when compared with a healthy control group, all glioma patients were found to be cognitively impaired, with 52% of the NSCLC group also falling into this category. By comparison with the NSCLC patients, the glioma patients were specifically impaired on some aspects of perception, memory and attentional and executive functioning. However, after correcting for differences in motor and visual functioning of the glioma group, significant impairment in memory function only was observed in the glioma group. This suggests that the cognitive differences between the two patient groups are largely the result of impaired vision and/or motor abilities in the glioma patients. However, only 37% of the high-grade glioma patient sample had a diagnosis of glioblastoma multiforme, yet this type of tumour accounts for the majority of high-grade glioma diagnoses. Therefore, given that GBM patients tend to present with more severe symptoms, the findings of this study may actually underestimate the impact of high-grade gliomas on cognitive functioning.

Bosma et al. (2007) aimed to map the neurocognitive functioning of patients with high-grade gliomas throughout the course of their disease, using data obtained as part of the aforementioned study by Klein et al. (2001). These authors evaluated the cognitive function of newly diagnosed high-grade glioma patients and also specifically examined the effect of tumour recurrence on cognition. Patients completed various neurocognitive tests at a baseline session after surgery but before radiotherapy and then again 8 months and 16 months later, where possible. The

neurocognitive tests that comprised the test battery provided measures of information processing and psychomotor function (Letter-Digit Substitution Test), attention (Stroop Colour-Word Test and Categoric Word Fluency), verbal memory (Visual Verbal Learning Test), working memory (Working Memory Task) and executive functioning (Concept Shifting Test). Premorbid functioning was also controlled for by means of the Dutch version of the New Adult Reading Test (DART). The HGG patients were divided into three groups for analysis: i) patients tested at baseline (n = 36), ii) patients tested at baseline and 8 month follow-up only (n = 14) and iii) patients tested at baseline, 8 month and 16 month follow-up (n = 18). Bosma et al. (2007) found no statistically significant differences in neurocognitive performance at baseline between the three patients groups with different follow-up lengths. However, patients who had tumour recurrence during follow-up had significantly poorer information-processing capacity and psychomotor speed (as measured by the Letter-Digit Substitution Test) at baseline testing than did those patients without recurrence. Overall, between baseline and subsequent testing at 8 months, there was a significant deterioration in information-processing, psychomotor speed, and attentional functioning (as measured by the Stroop Test and Categoric Word Fluency). Moreover, patients with tumour recurrence at 8 month follow-up performed significantly less well on tests of information processing, psychomotor speed and executive functioning than those without recurrence. However, after correcting for differences in drug use this effect was no longer observed. At 16 month follow-up, patients with tumour recurrence had significantly worse psychomotor speed than did those without progression, and this difference could not be accounted for by differences in drug use, since the use of antiepileptic medication and dexamethasone was equally divided between the two groups. Bosma et al. (2007) therefore conclude that there is a marked decline in cognitive functioning in high-grade glioma patients throughout the course of their disease. The study also confirms the importance of cognitive deficit as a potentially negative prognostic factor in glioma patients (Curran et al., 1993) since patients who went on to develop earlier tumour recurrence performed significantly less well on psychomotor speed measures than those who did not experience tumour progression during the study.

Therefore, although the exact role of a high-grade glioma as a cause of cognitive impairment has not been fully determined due to methodological difficulties associated with studying groups of brain tumour patients, it would appear that this group of patients do experience a wide range of difficulties in different domains of cognition.

1.3.2 Low-grade gliomas

Cognitive deficits in patients with a diagnosis of low-grade glioma have also been studied, although perhaps in less detail than their high-grade counterparts. Measuring and assessing the cognitive status of low-grade glioma patients is particularly important given that many patients have prolonged survival times (Correa et al., 2007). The utility of carrying out detailed neuropsychological assessments as opposed to brief assessment, or neurological examination only in low-grade glioma patients has been emphasised in a study that found moderate or severe cognitive impairment in more than half of the patients in a cohort of 35 individuals with a low-grade glioma, despite the fact that impairment was not detected by neurological examination and was not always reported by the patients themselves (Pahlson et al., 2003).

Given the scarcity of literature detailing the neuropsychological presentation of low-grade glioma patients, particularly those with tumours of the astrocytic subtype, Sweet et al. (1994) report a case series of six patients with low-grade astrocytoma who were not having active treatment at the time of enrolment into the study. Each patient completed assessments of intellectual function, cognitive efficiency, sustained attention, learning and memory, sensory/perceptual and visual/spatial abilities and language function. The authors report results for each participant separately, however, overall, it was found that each patient exhibited slight dysfunction, when compared with standardised scores, on some tests of cognition. The observed dysfunction was not nearly as significant as that demonstrated in studies of high-grade glioma patients and the mild deficits observed in this case series of patients were varied and did not routinely involve similar aspects of

cognitive functioning. Sweet et al. (1994) discuss possible reasons for which this group of patients perform relatively well neuropsychologically. They propose that, since low-grade astrocytomas tend to infiltrate widely among normal brain cells and do not destroy brain tissue in the manner that is characteristic of a more solid mass of tumour, this may explain why cognitive abilities appear to remain grossly intact in a high percentage of these patients. However, there are a number of important limitations associated with the methodology employed in this study, notwithstanding the extremely small sample size. Impairment on the cognitive tests was determined by comparison with normative data and therefore no control group was recruited and matched on important variables such as age, sex, number of years of education and premorbid ability, for example. Since no attempt was made to control for prior ability, the slight dysfunction observed in this study may be the result of lower premorbid ability in this small number of patients. Conversely, comparing the patients to a well-matched control group may have revealed that the impairment was actually more extensive than it appears. Moreover, since only six patients were recruited into this study, no definitive conclusion can be made with regard to the extent, and indeed the very presence of, cognitive impairment that results from a low-grade glioma.

In a further study examining the role of the tumour as a cause of cognitive deficit, Reijneveld et al. (2001) recruited two groups of low-grade glioma patients. They compared functional status, quality of life and the cognitive ability of patients who were suspected of having a low-grade glioma with that of those who had a proven diagnosis, having undergone early surgery. The suspected low-grade glioma group had a radiological diagnosis but surgical treatment had been deferred until the tumour showed signs of progression on CT or MRI scanning. The study aimed to determine whether uncertainty about diagnosis had any negative impact on cognitive ability and/or quality of life, as it has been shown to in other patient groups. If this proved to be the case, it would suggest that early intervention in the form of surgery in these patients may be a better policy than a "watch and wait" one, given the controversy surrounding the best way to manage these patients (Whittle, 2004). Twenty-four patients with a suspected low-grade glioma were recruited into the

study and completed various cognitive measures and an assessment of quality of life. The performance of the suspected low-grade glioma group was compared with data from a matched healthy control group and also with data obtained from patients who had undergone surgery to confirm the diagnosis of low-grade glioma ('proven'). It was found that both low-grade glioma groups (suspected and proven) had significantly poorer quality of life scores, as assessed by the self-report Short-Form Health Survey and Brain Cancer Module 20, than the healthy control group. Cognitive function was assessed using a battery of tests that included the Visual Verbal Learning Test, Working Memory Task, Letter-Digit Substitution Test, Categoric Word Fluency, Concept Shifting Tests, Stroop Colour Word Test and Working Memory Task. Both the suspected and proven LGG groups scored significantly worse on these measures by comparison with the healthy controls. However, patients with a suspected low-grade glioma performed significantly better overall than did those with a proven low-grade glioma. Specifically, patients with a proven glioma performed significantly less well on the psychomotor processing assessment than did the patients with a suspected glioma. Thus, this study suggests that impaired cognition is common to low-grade glioma patients and, given that impairments are evident in patients who have not yet had any intervention (i.e. surgery), the tumour itself likely plays some role in worsening cognition in this patient group. However, surgery may exacerbate any deficits that are present prior to intervention. The groups were well-matched in terms of demographic and clinical characteristics, including scores on commonly-used functional scales such as the Barthel Index and Karnofsky Performance Scale and had similar sized tumours (as measured by tumour diameter). However, the patients enrolled into this study all presented with epileptic seizures and were presumably treated with anti-epileptic drugs, although this is not documented. Given the aforementioned study by Bosma et al. (2007) who proposed that anti-epileptic drugs may play some effect on cognition, the cognitive deficits displayed by the low-grade glioma patients in this study may in fact be the result of prescribed drug use and not due to the presence of the tumour itself.

However, further evidence that proposes the presence of tumour-related cognitive impairment in patients with a low-grade glioma comes from a study by Ek et al. (2005) who retrospectively assessed cognitive function in a sample of low-grade glioma patients. Twenty-four patients with either stable disease or slowly growing tumours were recruited and each patient was administered a detailed neuropsychological examination. Cognitive measures that assess different domains were chosen to comprise a battery that could be completed in 1½ – 2 hours. The tests used included the Symbol Digit Modalities Test, Rey Auditory Verbal Learning Test various subtests from the Swedish version of the Wechsler Adult Intelligence Scale and Judgement of Line Orientation. In order to determine any possible patterns of cognitive dysfunction that may be present in the tumour group, the authors rated the patients according to the number of tests scores that were below normal limits (after correcting for age and education). A pattern of variation was found that resulted in separation of the patients into three subgroups with significant differences between the mean z-score. The first of these groups comprised patients who had a 'global' cognitive dysfunction and exhibited impairments in most domains that were assessed. The second group had slowed information processing-speed and also were impaired on between two and five of the other test domains. Patients in the third group generally performed well in most of the tests, including those of processing-speed. Additionally, it is reported that the majority of patients classified into the third group on the basis of good cognitive performance were found to have the most favourable prognosis based on the histopathological subtype of low-grade glioma. This finding that cognitive function is more impaired in patients with a low-grade glioma with a less favourable prognosis (i.e. more aggressive subtypes) is in keeping with the belief that more aggressive tumours (in particular high-grade gliomas) result in more pronounced cognitive deficits. However, the surprising finding of this study was the extreme variation in cognitive ability in this group of patients, with some patients showing severe disturbance yet others scoring within 'normal' limits. However, given that 17 of the patients recruited had received radiotherapy and a further 6 of those 17 also received chemotherapy, the differences in cognition may actually reflect differences in treatment, irrespective of the presence of the tumour itself. The authors did not analyse the potential effects of these treatments on

cognitive test scores due to the small number of patients and this makes it difficult to determine the effect of the tumour itself on cognitive function.

The majority of studies that address cognition in specific histological brain tumour groups are therefore unable to elucidate the exact contribution of the tumour as a cause of cognitive dysfunction due to treatment-related heterogeneity in the samples of patients recruited. Goldstein et al. (2003) highlight this issue and attempted to overcome this problem by recruiting patients prior to radiotherapy or chemotherapy. They specifically investigated the effect of low-grade tumours on attention function in a sample of 55 patients, 41 of whom were described as having 'superficial' tumours and 14 of whom had what was referred to as a 'deep' tumour. Neuropsychological testing took place 6 weeks following surgical excision or biopsy of the tumour, if this had taken place. Each patient completed a battery of standardised tests that were designed to measure specific domains of attention. The tests included the Auditory Selection Test, Bells Test, Symbol Digit Modality Test (Oral Version), Wisconsin Card Sorting Test, Digit Span Test, Visual Memory Span Test, and the Paced Auditory Serial Addition Test. The low-grade patients were compared to a normal control group, matched for age (n = 63). The normal control group had better overall performance than the low-grade brain tumour patients suggesting that these tumours have a negative impact on several different measures of attentional abilities. Moreover, the "deep" tumour group performed similarly to the other tumour patients on all the attentional measures, other than the Digits Forward task, on which they performed significantly less well. This study therefore suggests that low-grade brain tumours are particularly associated with impaired attentional ability, irrespective of the location of the tumour. However, this study differs from the aforementioned studies examining low-grade *glioma* patients in that it recruited any patients with low-grade brain *tumours*. Therefore, patients with meningiomas and pituitary adenomas were also included in the study. Given that these tumours are generally less infiltrative than gliotic brain tumours; the inclusion of patients with these types of tumour limits the conclusions that can be made regarding the specific effects of low-grade gliomas. Moreover, the control group used in this study was somewhat inadequate given that they were not matched with

the brain tumour patients on key variables such as premorbid ability or sex and this also questions the conclusions drawn in this study. Moreover, a more reliable assessment of the role of the tumour itself in contributing to cognitive impairment could have been achieved had the patients been recruited prior to surgery, instead of post-operatively but prior to radiation therapy.

1.3.3 High- and Low-Grade Gliomas – Conclusions

Therefore, there are a number of important methodological issues and flaws that are frequently associated with studies that attempt to determine the relative contribution of high-grade and low-grade gliomas themselves to cognitive impairment. However evidence consistently suggests that cognitive functioning in a variety of domains is impaired in many patients with both high and low-grade gliomas. Moreover, the impairments observed in high-grade glioma patients tend to be more severe than those measured in low-grade glioma patients. However, few studies specifically compare the two tumour groups explicitly. There exists a great variety in terms of the extent of impairment reported by different studies and this could be due to the presence of confounding variables such as differences in treatments between patients groups and methodological differences between studies, with different tests used and different means of classifying 'impairment'.

1.3.4 Other tumour types

Although the majority of studies that examine tumour-related impairment in cognitive function have focused on the effects of high and low-grade gliomas, some studies have addressed the specific cognitive effects of brain metastases and meningiomas.

Although it has been reported that 10 – 20% of all cancer patients will develop brain metastases, studies assessing quality of life and neurocognitive function in patients with metastatic brain tumours have rarely been carried out (Herman et al., 2003). Thus, a pilot study was conducted to address the feasibility of administering a

detailed neurocognitive test battery to patients with brain metastases, given the limitations associated with assessment using the Mini-Mental State Examination in this patient group. The MMSE, although brief and relatively non-demanding, is less sensitive to mild impairment and is not really suitable for repeated assessments due to practice effects. Herman et al. (2003) recruited 30 adult patients with any newly diagnosed, existing or recurrent brain metastases and administered a five-part neurocognitive test battery to each patient. The tests that comprised the battery were all widely-used, standardised measures of cognition and included the Hopkins Verbal Learning Tests, Trail Making Parts A and B and Controlled Oral Word Association (COWA, a measure of verbal fluency). The results of this preliminary investigation showed that the group of patients with a brain metastasis were impaired on measures of recall and delayed memory. Impairment was defined as a score more than 1.5 standard deviations from the mean of normal age-matched subjects for the Hopkins Verbal Learning Test or as a score below the tenth percentile for the normative population for the COWA and Trail Making tests. Specifically, 76% of patients exhibited significantly impaired recall and 70% displayed significant impairment in delayed memory. Although some patients were classified as impaired on other tests including Trail Making Tests A and B and the Controlled Oral Word Association test, the proportion who displayed impairment was much smaller than that seen for the tests of memory. The authors conclude that patients with brain metastases can tolerate a 30 minute neurocognitive test battery and that such measures should be added prospectively to clinical trials as an outcome measure in this patient group. However, it is important to consider the fact that the 'control' group data in this study was derived from normative data from age-matched subjects. Thus, no attempt to control for variables such as education, premorbid ability, sex and the potential effects on increased emotional distress on test performance was made, limiting the conclusions that can be drawn regarding the cognitive status of this patient group. However, the study does highlight the potential utility and importance of carrying out further studies that assess cognition in patients with metastatic brain tumours.

In a similar study that tested the feasibility of carrying out neurocognitive measures and quality of life assessments in a group of patients with brain metastases, Regine et

al. (2004) report high compliance rates, with the majority of participants (> 90%) yielding valid data on a number of different cognitive assessments. The authors do not give any information detailing the specific tests administered or the actual scores obtained and therefore do not report the presence or otherwise of any impairments in this group but the study again paves the way for future research in this under-studied area.

The results of cognitive assessments of glioma patients cannot equally be applied to patients with meningiomas for a number of reasons. Namely, gliomas tend to be more invasive, and high-grade tumours in particular grow much more rapidly than meningiomas which tend to be benign, slow-growing and non-infiltrative. Therefore meningiomas could reasonably be expected to have different and potentially fewer effects on cognition. However, compared with gliomas, the cognitive effects of meningiomas have received far less attention in terms of their effect on cognitive functioning. A number of case studies have frequently reported disturbances in memory, concentration and orientation in patients with frontal meningiomas in particular (Avery, 1971). However, although a small number of studies have systematically examined cognition in larger cohorts of patients with meningiomas, many report only the incidence of neurological disturbance on admission, prior to any intervention being made, in addition to detailing postoperative functional outcome and prognosis (Feder et al., 1989, Simoca et al., 1994, Torres et al., 2003). Neurological symptoms frequently reported to be present in meningioma patients include aphasia, confusion and disorientation and, according to these studies, these symptoms were present in between 25% and 62% of meningioma patients during the course of their disease. However, no formal cognitive testing is reported by these authors.

In one of the few studies that have systematically assessed cognition in this patient group, Tucha et al. (2003), prospectively examined pre- and post-operative functioning in patients with frontal meningiomas, giving a detailed account of the cognitive deficits in this group. Fifty-four such patients were recruited prior to surgery and completed a detailed battery of tests assessing aspects of memory,

attention and executive functions both pre and post-operatively. A group of 54 healthy volunteers were administered the tests on two separate occasions to serve as a control group. The two groups were well-matched in terms of age, sex, handedness, intelligence, education level and the duration between first and second testing session. Tests included in the battery were Digit Span, Logical Memory, Visual Reproduction, Divided Attention Task, Trail Making Test Parts A and B, the Five Point Test and Complex Figure Test. Analyses detailing the effects of surgery on function are discussed in due course in this chapter (see section 1.5.1). Prior to surgery, significant differences in the performance of the meningioma group by comparison with the healthy control group were found. In particular, there was a significant preoperative impairment in working memory, fluency functions, tonic alertness, processing speed, shifting, divided attention, and flexibility in the meningioma patient group. However, no differences were observed in intellectual function, memory or visuoconstructive abilities. Therefore, this study suggests that the presence of the meningioma itself may cause significant impairments in several specific cognitive domains. It is one of few to detail cognition in brain tumour patients prior to *any* treatment (i.e. at presentation, before any surgical intervention) and therefore gives greater insight into the role of the tumour itself as a cause of cognitive impairment.

1.3.5 Tumour Type and Cognition: Conclusions

There is little doubt that many patients with brain tumours suffer cognitive impairment and studies have shown a general trend towards these impairments being more pronounced in patients with more aggressive tumours (i.e. high-grade gliomas). However, deficits have consistently been recorded in studies of patients with low-grade gliomas and also to some extent in patients with other, less common tumour types. However, further research into the cognitive function of meningioma and metastasis patients, in particular, is required.

Although a number of factors are likely to play some role in contributing to these impairments, it does appear that the tumour itself causes cognitive deficits to some

extent. Screening for cognitive impairment prior to any treatment is particularly important as it not only allows for the most effective interventions to be offered to the patient but also provides a baseline level of function from which to assess the success of any subsequent treatment.

1.4 Brain tumour location and cognition

The *location* of the tumour within the brain has also received attention as a possible cause of cognitive impairment in neuro-oncological patients with studies aiming to determine whether tumours located within specific brain areas result in impairment in specific cognitive domains.

Early studies regarding the relationship between tumour location and cognitive impairment yielded mixed results, with some studies reporting poor verbal performance in patients with left-sided compared with right-sided tumours (Hom and Reitan, 1984), yet others reporting no such differences (Haaland and Delaney, 1981). Subsequently, Scheibel et al. (1996) used age-corrected test scores to examine the influence of tumour lateralisation, amongst other variables, on cognitive function in a large cohort of glioma patients. Two-hundred and forty-five glioma patients were recruited, the majority of whom had a diagnosis of glioblastoma multiforme. Each patient completed a detailed neuropsychological test battery and it was found that performance on several of the measures was dependent upon lesion lateralisation (i.e. whether the tumour was located in the left or right hemisphere of the brain). Specifically, patients with left hemisphere tumours had lower scores on measures of language, verbal learning and verbal intelligence, whereas patients with right-sided lesions had greater difficulty with visual-perceptual skills, as measured by a test of facial recognition. Although the study is limited by its retrospective nature, these findings are of interest since the results of previous studies have not always found a significant relationship between lesion site and type of cognitive impairment.

Subsequent studies in this area have reported, perhaps unsurprisingly, that patients with right-sided tumours score significantly worse than patients with left-sided

tumours on tasks that rely heavily on functions carried out by the right-hemisphere, such as the line bisection test and facial recognition (Klein et al., 2001). Additionally, left-sided tumours were found to have a deleterious effect on both attention and executive function, and this was reflected by longer scanning times in the Concept Shifting Test A, increased susceptibility to interference as exhibited by poorer Stroop Test performance and reduced fluency ability. Thus, this study provides further evidence to suggest that the hemispheric location of the tumour appears to have an effect on specific aspects of cognition.

Patients with a glioma located in the dominant hemisphere of the brain tend to have more cognitive deficits than do those with a non-dominant hemisphere tumour, a finding that was confirmed by Hahn et al. (2003) whose study is described in detail previously (see section 1.3.1). Their study also examined the effect of lesion lateralisation on cognitive test performance. The performance of the right and left hemisphere tumour groups were compared and a significant relationship between hemispheric tumour site and neuropsychological test performance was found. The participants in this study were left hemisphere dominant and univariate analyses revealed that these patients with left hemisphere tumours (i.e. in the dominant hemisphere) performed significantly less well than patients with tumours in the non-dominant right hemisphere on the COWA test (a measure of verbal fluency), 2 and 7 Continuous Performance Test (attention and reaction time), Victoria Stroop Latencies (cognitive flexibility) and on the delayed free recall, recognition, and consistency of long-term retrieval from the Levin Selective Reminding Test. The authors thus conclude that left-sided tumours have a negative effect on memory, verbal fluency and verbal learning functions. This finding also supports conclusions made by Hom and Reitan (1984) who report poorer verbal functioning of patients with tumours located in the left hemisphere and also correlates with the findings of the aforementioned studies (Scheibel et al., 1996, Klein et al., 2001).

Thus, evidence exists to suggest that tumour location has an effect on several aspects of cognitive function with dominant hemisphere lesions resulting in greater impairments compared to tumours located in the non-dominant hemisphere.

However, given that many patients, particularly those with a low-grade glioma, frequently experience non-specific cognitive slowing and diffuse deficits that cannot simply be explained by tumour location, Bartolomei et al. (2006b) carried out an investigation to test the hypothesis that brain tumours interfere with normal brain function through the disruption of functional connectivity of brain networks. Previous studies have shown that cognitive function in healthy participants relies upon interactions between multiple discrete neural networks in the brain and complex functional interactions between different brain areas have been found using magnetoencephalography (MEG) or electroencephalography (EEG) signal analysis, even when the participant has not been asked to complete any cognitive task (Stam et al., 2002). Thus, the study aimed to determine whether there is a loss of functional connectivity in the brains of patients with low-grade gliomas, and whether any loss of connectivity is localised solely within the region in which the tumour is located, or whether it extends to different brain regions distant from the tumour itself. Seventeen histologically confirmed low-grade glioma patients and 15 healthy control subjects had MEG recordings analysed. Bartolomei et al. (2006b) found a loss of functional connectivity in the brain tumour group compared with the control group, and this extended to multiple brain regions. Therefore, brain regions beyond the margins of the tumour itself were disrupted. The disruption was more evident in patients with tumours located within the left hemisphere. Thus, functional networks in the brain that are required for cognitive test performance were disturbed in various areas of the brain, including areas that appear to be 'normal' on MRI scanning. This preliminary study therefore proposes that disruption of functional networks throughout the brain as a result of the presence of a tumour may explain the observed global cognitive impairment (e.g. in psychomotor and executive function domains) in low-grade glioma patients cannot be explained by focal tumour-related effects. However, given that all the patients in the study had already undergone surgery to obtain a diagnosis of low-grade glioma, the potential role of surgery in disrupting functional networks cannot be distinguished from the effect of the tumour itself. Additionally, no cognitive testing was carried out in order to determine the presence and/or the extent of cognitive impairment in the patients recruited into the study.

Thus, Bosma et al. (2008) carried out a further investigation using MEG analysis to relate functional connectivity directly to neurocognitive function in 17 low-grade glioma patients and 17 age, sex and education-matched healthy controls. MEG recordings were taken for each participant, each of whom then completed a detailed cognitive test battery that included measures of psychomotor function (Letter-Digit Substitution Test); executive function, attention and visual scanning (Concept Shifting Test); attention, mental speed and mental control (Stroop Colour Word Test); verbal learning, organisation and memory (Visual Verbal Learning Test); selective attention, mental concentration and information processing (Memory Comparison Test) and frontal dysfunction and flexibility (Categoric Word Fluency). The low-grade glioma patients were found to be impaired on measures of psychomotor function, working memory, information processing speed and attention by comparison with the matched control group. The findings of the aforementioned study (Bartolomei et al., 2006b) was confirmed by this follow-up study. The low-grade glioma patients were found to have significant differences in resting-state functional connectivity compared with the healthy control group. This again suggests that the tumour causes disruption to functional networks throughout the brain. Moreover, loss of functional connectivity in the delta, theta and lower and upper gamma bands and neurocognitive function were significantly correlated in the low-grade glioma patient group, thus suggesting that cognitive impairment, specifically in the domains of attention and information processing, in brain tumour patients may arise in part as a result of disturbed functional connectivity in these specific wave bands.

Thus, evidence exists to suggest that the location of the tumour, in particular the hemisphere in which it is located, may cause specific cognitive impairment. However, recent research has shown that the more global executive function and information processing impairments that are particularly evident in patients with low-grade gliomas may arise as a result of disruption to the discrete functional networks in the brain that are involved with performing cognitive tasks.

1.5 Treatment effects on cognition

As has been shown, although the causes of cognitive impairment in brain tumour patients are not fully understood, it is likely that tumour type and location within the brain both contribute to some degree. Studies have also addressed the role of treatment-related variables including surgery, radiotherapy and chemotherapy, steroid therapy and drugs such as anti-epileptic medications that are commonly prescribed to patients who suffer seizures. Examining the treatment-related cognitive sequelae is of great importance in neuro-oncology, since the majority of patients have a poor prognosis. Therefore a treatment that has little or no effect in terms of improving survival may still be beneficial if it improves or even maintains cognitive function since impaired cognition can result in poor quality of life. Conversely, the utility of a therapy that only modestly prolongs survival may be questionable if it has a significantly detrimental effect on cognitive function.

1.5.1 Surgery and cognitive function

The first-line treatment for a brain tumour is most often surgery in the form of resection or biopsy to obtain a histological diagnosis, and in some cases, to alleviate symptoms of increased intracranial pressure. Surgical intervention is not without risk and tumour biopsy or resection may cause or exacerbate existing deficits due to damage of normal brain tissue that surrounds the tumour (Taphoorn and Klein, 2004). A number of studies have addressed the potential relationship between surgery and cognitive impairment.

Vecht et al. (1990) assessed the effect of the extent of surgery on the neurological function more generally, as opposed to cognitive function, of high-grade glioma patients. A retrospective analysis of 66 patients with anaplastic astrocytoma and 177 patients with GBM indicated that more extensive surgery as opposed to limited resection did not have any negative effect on post-operative neurological function. Although this early study did not formally assess cognitive function, it does suggest

that extensive surgery to remove a high-grade glioma does not impact negatively on neurological status, at least in the initial post-operative period.

Whittle et al. (1998) looked specifically at the effects of resective surgery for left-sided tumours on language functions, since language impairment is a common symptom of a left hemisphere lesion. Forty consecutive patients were recruited and completed the Western Aphasia battery (WAB) and the Boston Naming Test (BNT) preoperatively and postoperatively, prior to discharge. The WAB is a group of subtests assessing aspects of auditory and reading comprehension in addition to oral and written expression. The BNT is a test of confrontation-naming that yields a score out of 60 with score of 48 or less taken to represent impairment. It was found in the study that language function, as measured by the WAB and BNT, actually improved in 92% of patients who had some degree of dysphasia prior to surgery. This suggests that, in the case of language dysfunction, surgery may actually relieve pre-existing impairment. Moreover, of those patients who scored within normal limits on the language assessments prior to surgery and therefore were not deemed to be dysphasic, none showed any significant postoperative change. Thus, resective surgery for left-sided brain tumours does not appear to negatively impact language function in the initial pre-operative period and may actually offer some benefit in improving speech in those patients who were dysphasic pre-operatively. However, since no control group was recruited into this study, the potential effects of practice on the tasks administered are unclear and may account for some of the improvement observed in such a high proportion of patients. Further studies should be carried out in this area, to support the conclusions made by Whittle et al. (1998).

However, in contrast with the idea that surgery tends to alleviate cognitive dysfunction, the study by Reijneveld et al. (2001) described previously (see section 1.3.2), reports that patients with a histologically proven low-grade glioma (i.e. those who had undergone surgery to confirm the diagnosis) performed significantly less well on measures of psychomotor function and health-related quality of life as assessed by the MOS SF-36 Short-Form Health Survey, than did a case-matched group of patients with a low-grade glioma diagnosed on the basis of MRI or CT

scanning alone. This would therefore suggest that, in this group of low-grade glioma patients, surgery caused, or at least exacerbated the severity of existing cognitive deficits. However, as the authors highlight, it is not possible to confirm whether the decreased cognitive functioning of patients with proven low-grade gliomas should be attributed to the surgical procedure, or whether other variables may account for the differences. These include potential differences in anxiety levels between the groups, given that mood was not measured, in addition to the potential for selection bias since those patients for whom surgery was deferred were likely to have fewer impairments than those who had earlier surgical intervention. No information regarding the pre-operative neuropsychological status of these patients was available, and this limits any conclusions regarding the contribution of surgery to cognitive dysfunction in the study.

Teixidor et al. (2006) provide further evidence to suggest that brain tumour surgery, at least in low-grade glioma patients, may cause a worsening of existing deficits, at least in the initial pre-operative period. Verbal working memory was examined before and after surgery in a group of patients who had awake surgery to resect a low-grade tumour. Twenty-three such patients were recruited into the study and verbal working memory was assessed before and immediately after surgery. A further 8 patients had the assessment for a third time, 3 months post-surgery. Results indicated that, pre-operatively, 91% of the patients tested had impairment in verbal working memory. This provides further evidence to suggest that cognitive impairment may be a result of the tumour itself in these patients. Furthermore, in the immediate post-operative period, 96% of the patients had worsening of verbal working memory, with the authors reporting that 22 of the 23 patients had lower scores on the neurocognitive assessments following resection of the tumour. In five of the nine tests administered, the difference between pre- and post-operative scores was found to be statistically significant in the group of low-grade glioma patients. However, of the 8 patients studied 3 months later, 5 recovered their preoperative scores and a further 3 actually improved significantly on the assessment of verbal working memory. This again suggests that cognitive impairments, like physical neurological impairments that also commonly occur after awake surgery (Duffau et

al., 2003) are mainly transient and scores may actually exceed preoperative ones 3 months after surgery.

In the aforementioned study of frontal lobe meningioma patients, 54 patients were assessed neuropsychologically both pre and post-operatively (Tucha et al., 2003). The tests included in the battery are detailed in section 1.3.4 of this chapter. There were no significant differences between the meningioma patient's pre- and post-operative scores on the tests of memory (with the exception of working memory), visuoconstructive abilities or executive functions. Several attentional functions, including tonic alertness, divided attention, flexibility and shifting were actually found to improve post-operatively in this patient group, suggesting that surgical removal of the tumour did not have any negative impact on the cognition. However, given that the same standardised cognitive measures were administered to patients pre and post-operatively, practice effects may explain the improvements observed to some degree. The authors suggest that the effect of practice is likely to have had been minimal on the attentional measures since the tests used were generally designed to be used for repeated assessment. However, to enhance the robustness of their findings, the test performance of a group of healthy volunteers who were tested on two separate occasions were analysed. There was found to be a significant improvement in immediate and delayed recall of verbal information, with no other significant changes found on the other tests. Thus, this study supports the proposal that surgery for a brain tumour does not tend to negatively impact cognition, at least in meningioma patients.

Therefore, evidence to suggest that surgery may play some role as a cause of cognitive deficits in brain tumour patients is somewhat conflicting, with some studies proposing that it has little or no negative effect and may actually alleviate existing deficits to some degree, yet others suggest that surgery results in deterioration in cognitive function. Any impairments following surgery may however be mainly transient, and this emphasises the importance of carrying out extended follow-up assessments in order to determine whether any specific long-term effects of surgical intervention exist.

1.5.2 Radiotherapy and cognitive function

Following surgical intervention, dependent upon the age and functional status of the patient, those with malignant gliomas (WHO grade III and IV) are likely to be offered immediate radiotherapy. The effects of radiotherapy on the brain, and in particular on cognitive function, have been widely studied with varying results. Until the mid 1980s, brain tumour patients receiving radiotherapy were given standard whole-brain radiotherapy (WBRT). There is little doubt that WBRT had significant effects on cognitive function with Gregor et al. (1996) showing that short-term memory was particularly affected in those patients who survived for more than 5 years following treatment. However, radiotherapy treatment protocols have changed significantly over the past 20 years with patients now commonly offered partial brain irradiation that is targeted specifically towards the brain areas invaded by tumour, using smaller fraction sizes. Studies have generally found this treatment protocol to have less of an impact on cognition (Taphoorn and Klein, 2004).

The effects of radiation therapy on the cognitive function of patients with low-grade gliomas in particular have received a great deal of attention. This is due to the longer survival of these patients and well-documented controversy surrounding the question of whether to offer patients with a low-grade glioma immediate radiotherapy, or whether to withhold treatment until such a time as the tumour progresses (Whittle, 2004). A randomised trial reported that although early radiotherapy after surgery lengthens the time to tumour progression, it has no effect on overall survival (van den Bent et al., 2005) and is now generally accepted that radiotherapy can be deferred in low-grade glioma patients who have a good performance status.

Armstrong et al. (2002) examined the effects of radiotherapy on both the cognitive and radiographic outcomes of 26 patients with low-grade brain tumours who were without risk factors for vascular damage (e.g. hypertension). The patients were tested on a yearly basis from baseline (6 weeks post-surgery, prior to radiotherapy) to 6 years. A test battery comprising a large number of standardised neuropsychological

tests that were designed to assess a number of different aspects of cognition was administered and scores on a total of 37 different tests were included in analyses. In contrast with previous studies that report between 20% and 80% of patients exhibit some degree of cognitive impairment between 2 and 20 years after receiving radiotherapy (North et al., 1990, Meyers et al., 2000) Armstrong et al. (2002) found no evidence of any general cognitive decline in these low-grade glioma patients 3 years after treatment. However, a late-delayed effect of radiotherapy was demonstrated by a decline in selective areas of cognition, namely in visual memory, that began 5 years following radiotherapy. In contrast, MRI scan images showed white matter atrophy from 6 months to 3 years post-treatment, with no further progression evident after this time. The study is limited by the high levels of attrition that occurred 2 years post-radiotherapy as it is likely that those patients who did not complete the full follow-up were those patients who had the most severe functional and/or cognitive deterioration. As such, the study may underestimate the potential effects of radiotherapy on cognitive function.

Torres et al. (2003) investigated the effect of modern (partial) brain radiation therapy on the cognitive functioning of low-grade brain tumour patients in a longitudinal study. Seventeen patients with low-grade gliomas were tested on a battery of neuropsychological assessments pre-radiotherapy (baseline) and at 3, 6, 12 and 24 months subsequently. The tests were chosen on the basis of their sensitivity to brain dysfunction, their emphasis on attention, processing speed and memory and their suitability for repeated administration. Twelve of the 17 patients had no tumour progression during the course of the study and these patients showed a slight improvement in verbal learning between baseline and 3 month follow-up that could likely be attributed to practice effects. There was no evidence of decline on any of the other measures in these patients. Between 3 and 6 months, 6 months and 1 year, and 1 and 2 years, there was no decline on any of the tests administered in this group of low-grade glioma patients without tumour progression. Therefore, no ‘early-delayed’ cognitive effect of radiotherapy was observed as described by previous studies (Vigliani et al., 1996). The three patients who were ‘progressors’ during the study showed a general decline in test performance between baseline and 1 year by

comparison with two patients who had surgery but not radiotherapy, who served as controls. Thus, this study found evidence to suggest that there is a favourable outcome in terms of cognitive function in the first 2 years following partial radiotherapy, at least in adult patients with low-grade tumours. However, there are numerous methodological criticisms that can be made of this study. Primarily, the sample size was very small, with only two patients serving as control patients. These 'controls' were not matched with the study group with respect to age, sex and education and the lack of randomisation of patients to radiotherapy and non-radiotherapy means that there was likely some selection bias surrounding which patients were selected for post-operative radiotherapy. The authors themselves suggest that these limitations should be considered when generalising the results to other patients, especially those out with the treatment centre in which the study was conducted.

Although the majority of studies that have examined the potential early and delayed effects of radiotherapy have focused on low-grade glioma patients, some studies have focused solely on radiotherapy-effects in high-grade glioma patient cohorts. Taylor et al. (1998) carried out a prospective study in which they administered the Folstein Mini Mental State Examination (MMSE) at baseline, 6, 12, 18 and 24 months to a group of high-grade glioma patients who were treated with combined radiotherapy and chemotherapy. Significant cognitive decline (described as a greater than 3 point decrease in baseline MMSE score) in patients without tumour progression, was observed in 11% (13/119) of the patients at 6 month evaluation, 6% (3/54) of the cohort at 12 months, 10% (3/30) at 18 months and 18% (4/22) at 24 months post-treatment. Those patients who had a significant decrease in MMSE score were more likely to be older and to have a poorer performance status. The authors conclude that there was no clear trend towards increasing cognitive impairment in this group of patients following radiotherapy and chemotherapy treatment. However, in addition to a high rate of attrition which may have biased the results, the use of the MMSE as an assessment of cognition in this study is a significant limitation. The MMSE has been validated for use with dementia patients but has not been validated for use in brain tumour patients. Meyers and Wefel (2003)

state that the MMSE lacks sensitivity when used to assess cognitive function in cancer patients in general. In particular, they highlight the fact that it is particularly unsuitable for use in radiotherapy trials since it does not assess the cognitive functions that are most likely to be affected by radiotherapy, such as learning and memory, processing speed, executive function and fine motor control.

Therefore, a number of studies have assessed the cognitive effects of radiotherapy, with the majority showing that focused brain irradiation has fewer and less severe cognitive deficits than whole brain radiotherapy. However, despite these treatment advances, there is still likely to be some degree of cognitive impairment following radiotherapy and this may not be evident until a number of years following treatment.

1.5.3 Prescribed Medication and Cognition

1.5.3.1 Antiepileptic drugs

Epileptic seizures are a common presenting symptom of a low-grade glioma with up to 80% of these patients presenting with seizure disorder (Whittle, 2004). The potential deleterious effect of antiepileptic drugs on cognitive function has been addressed in a small number of studies. In a study of 195 low-grade glioma patients that aimed to differentiate between tumour and treatment effects on cognition, it was found that patients on antiepileptic medication performed significantly less well on several attentional and executive function measures compared with patients who were not taking antiepileptic drugs. The patients who had been prescribed antiepileptic medication also had significantly reduced self-reported cognitive function (Klein et al., 2002). Seizures themselves may also contribute to cognitive decline. In a study of patients with refractory temporal lobe epilepsy it was found that those patients with a longer duration of epilepsy were most severely impaired on measures of psychometric intelligence, as measured by the full scale intelligence quotient of the Wechsler Adult Intelligence Scales (Jokeit and Ebner, 1999).

Therefore, Klein et al. (2003a) recruited 156 low-grade glioma patients with an epilepsy burden ranging from 'none' to 'severe' (based on seizure frequency and use of antiepileptic drugs) and compared the cognitive ability of the LGG group with that of a matched group of healthy controls. The study aimed to determine the impact of both epilepsy itself and of antiepileptic drug treatment (AEDs) on cognitive functioning and health-related quality of life (HRQOL). The LGG group were divided into one of six groups dependent on their epilepsy burden, ranging from 1 - epilepsy free to 6 - epilepsy, more than 6 seizures in the previous year. Of the LGG patient group, 86% had epilepsy and had been prescribed AED therapy. A comprehensive test battery that measured the domains of information processing speed, psychomotor function, attentional functioning, verbal memory, working memory and executive functioning was completed by each patient. The Medical Outcomes Study (MOS) Short-Form Health Survey was also administered as a measure of health-related quality of life. The tests in the neurocognitive test battery included the Line Bisection Test, Facial Recognition Test, Judgement of Line Orientation Test, Letter-Digit Substitution Test, Visual Verbal Learning Test, Working Memory Task, Stroop Colour-Word Test, Categoric Word Fluency Task and Concept Shifting Test. After controlling for differences in age, sex and education, it was found that the LGG patient group had significantly lower scores than the healthy control groups on measures in all of the aforementioned domains. A higher epilepsy burden was associated with slower information processing speed, impaired psychomotor function, working memory capacity and executive functioning. However, there was no association between higher epilepsy burden and impaired attentional functioning or verbal memory capacity. Further analysis suggested that patients using antiepileptic drugs performed less well than those not using antiepileptic medications on all domains, with the exception of verbal memory. This suggests that AEDs have a specific, negative effect on cognitive function, yet the seizures themselves have little impact. Although the tumour and treatment characteristics varied between patients, there was no significant overall difference in terms of sex, educational level, histological diagnosis, tumour lateralisation or treatment (surgical intervention and/or radiotherapy) in each of the patient subgroups and this minimises the potential confounding effects of these variables. Thus, the

potentially deleterious effects of AEDs should be considered in studies addressing cognitive function in cohorts of brain tumour patients, and in particular in studies of low-grade glioma patients.

1.5.3.2 Steroids

Corticosteroids are prescribed to the majority of brain tumour patients at some point during the course of their disease, and are primarily used to alleviate tumour-related vasogenic oedema and the associated symptoms (Koehler, 1995). Dexamethasone is the most commonly used corticosteroid prescribed to patients with brain tumours (Taphoorn and Klein, 2004). A number of studies have reported impairments in several different cognitive domains associated with the use of dexamethasone, including a decrease in immediate and delayed free recall in healthy participants (Newcomer et al., 1994) and a positive correlation between severity of dementia and cortisol levels associated with use of dexamethasone in groups of patients with Alzheimer's disease (Balldin et al., 1983, Davis et al., 1986). Corticosteroids can also cause a steroid-induced psychosis, although this is a relatively rare and usually reversible complication (Kershner and Wang-Cheng, 1989). Thus, dexamethasone may play some role in causing cognitive impairment in brain tumour patients. However, it is more likely that steroid therapy in brain tumour patients works to alleviate cognitive impairment by reducing tumour-related oedema (Taphoorn and Klein, 2004).

1.6 Cognitive function as a prognostic indicator

Measurement of cognitive function in brain tumour patients has a potential utility as a measure of the efficacy of a particular treatment, with or without an increase in median survival time. More recently, studies have offered evidence for a further practical utility of cognitive testing in neuro-oncological patients: cognitive function, as measured by standardised tests, may serve as a prognostic indicator that could indicate tumour progression, even before any evidence of this can be seen on MRI

scanning. This could facilitate the application of earlier therapeutic interventions that could, in turn, result in improved outcome (Armstrong et al., 2003).

Therefore, a number of studies have examined the potential relationship between measures that directly reflect brain function (i.e. cognitive function) and the duration of survival, specifically in patients with a glioma for whom survival times are limited. Meyers et al. (2000) examined the potential for cognitive function to predict survival in patients with recurrent malignant brain tumours. Eighty patients with recurrent GBM or anaplastic astrocytoma were recruited, each of whom completed baseline testing before commencing any post-surgical treatment. Follow-up sessions took place on a monthly basis thereafter and 58 patients (73% of the total cohort) had at least one follow-up session. The test battery was selected specifically to minimize the effects of repeated administration and lasted around 40 minutes in total. Tests administered to each patient included Digit Span as a measure of attention, Digit Symbol (graphomotor speed), Hopkins Verbal Learning Test (memory), Controlled Oral Word Association (verbal fluency), Trail Making Test Part A (visual-motor scanning speed), Trail Making Test Part B (executive function), Grooved Pegboard (motor speed and dexterity), Functional Assessment of Cancer Therapy (quality of life) and Functional Independence Measure (activities of daily living). Univariate Cox analyses of the nine cognitive tests revealed that baseline performance on seven of the nine variables was significantly related to survival. After adjusting for several potential confounding variables including age, Karnofsky Performance Score (KPS), histology and diagnosis-to-test interval, performance on a measure of memory (Hopkins Verbal Learning test) remained significantly related to survival. Multivariate analyses revealed that the clinical variables, histology (i.e. GBM or anaplastic astrocytoma) and number of tumour recurrences were significantly related to survival. The extent of resection, age, KPS, time since diagnosis and the number of previous surgeries were not significantly related to survival. However, better performance on three of the cognitive tests – Digit Span, Digit Symbol and Hopkins Verbal Learning Test – was significantly related to longer survival. Models incorporating scores on these three tests with the clinical variables accounted for 49% of the variance in survival, compared with 34% when only the clinical variables

were included. When the nine cognitive scores were added to the clinical variables, 53% of the variance in survival was accounted for. Thus, according to this study, a combination of tumour prognostic variables and assessments of cognitive function appear to more accurately predict survival than clinical variables alone. This again highlights the potential importance of carrying out cognitive assessments with neuro-oncological patients.

Klein et al. (2003b) recruited 68 consecutive, newly-diagnosed patients with histologically-confirmed high-grade gliomas (WHO Grade III and IV) who were to receive radiotherapy, as part of a longitudinal study into cognitive functioning and health-related quality of life in high-grade glioma patients (Klein et al., 2001). The methodology of the study is described in section 1.3.1 of this chapter. In brief, cognitive, performance and functional status was assessed and the cognitive test battery included assessments of perception, memory, attention and executive function. At the time of analysis, 61 of the 68 patients included in the study had died and 7 patients were 'censored' and not included in the analysis due to a lack of survival data, although it unclear whether all 7 censored patients were still alive or whether some may have been lost to follow-up. The median survival of patients who had cognitive impairment was significantly shorter (6.8 months) than that of those who had no cognitive deficit (13.6 months). However, when age, tumour grade, lateralisation, size, neurosurgical intervention (biopsy or resection), performance status and cognitive status were all included in a multivariate Cox proportional hazards model, only older age and higher tumour grade were independently related to poorer survival. Although cognitive status was not found to be independently predictive of survival in this combined model, the presence of cognitive deficit at the time of surgery in older patients with WHO Grade IV glioma (i.e. patients older than 50 years; n = 42), but not in the other patient groups, was associated with significantly poorer survival. Therefore, this study suggests that, in older patients with GBM, cognitive functioning may have some prognostic value, although larger cohorts of patients may be required to confirm this finding. Again, this highlights the utility of cognitive testing, not only to provide important clinical information that can inform suitable interventions, but as a possible prognostic indicator, at least in older

patients with GBM. However, no attempt was made to measure premorbid ability in this study and a relatively small number of patients were enrolled which may question the validity of this finding. Nevertheless, this study paves the way for future research into this area that should include larger cohorts of patients with a greater age range.

Meyers and Hess (2003) have shown that cognitive deterioration may actually be the first sign of tumour progression, in the absence of radiological evidence of progression on CT or MRI scanning. Fifty-six patients with recurrent glioblastoma or anaplastic astrocytomas completed assessments of cognitive function, quality of life and ability to carry out activities of daily living prior to treatment and at regular intervals forthwith. Specifically, the test battery included measures of attention span (Digit Span test); graphomotor speed (Digit Symbol Coding); memory (Hopkins Verbal Learning Test); verbal fluency (Controlled Oral Word Association Test); visual-motor scanning speed and executive function (Trail Making Test Parts A and B) and motor speed and dexterity (Grooved Pegboard). Quality of life was assessed by the Functional Assessment of Cancer Therapy and activities of daily living were measured by the Functional Independence Measure. It was found that cognitive decline, as revealed by a battery of neurocognitive tests, occurred approximately 50% earlier (6 weeks) than MRI evidence of tumour progression. When only scores on the 3 tests most sensitive to cognitive decline were included in analyses, cognitive decline as measured by the Reliable Change Index (RCI) was still found to occur more than a month prior to MRI evidence of tumour progression. This decline was observed even although improvement in performance on the tests was expected due to practice effects. Thus, it was concluded that direct assessment of brain function by means of cognitive testing is more sensitive than MRI scanning evidence of time to progression and could predict tumour recurrence more than a month prior to MRI confirmation of this tumour progression. This finding again points to the importance of carrying out serial cognitive assessment in patients with brain tumours.

1.7 Issues associated with measurement of cognitive function in brain tumour patients

As has been shown, measurement of cognitive function in brain tumour patients at various times throughout the disease journey can convey numerous benefits in terms of assessing the patient's ability to carry out daily activities, offering targeted interventions to ameliorate the effects of such impairments and to guide treatment decisions. Detailed cognitive assessment can provide a measure of the efficacy of new and existing treatments and may provide a means of identifying tumour progression before structural evidence of this can be seen on an MRI scan. However, cognitive assessment of brain tumour patients can be particularly problematic and a number of different variables may threaten the validity of in-depth cognitive assessments.

Several demographic factors - namely age, gender and educational level are known to have a significant effect on cognitive test scores (Laursen, 1997). When analysing and comparing cognitive test scores between two or more groups of participants, it is important to not only control for the effect of age and gender during the analysis procedure but also to consider the premorbid ability of the patient. The National Adult Reading Test (NART) is commonly employed as an estimate of prior intelligence in studies of cognition in several different clinical populations. NART score is significantly correlated with measures of full-scale intelligence, and importantly, has been shown to provide an index of *prior*, as opposed to current, intellectual function in clinical populations (Crawford et al., 2001). However, given that the test involves reading aloud a list of words, performance is largely dependent upon intact speech function. Many patients with brain tumours will experience speech impairment to some degree during the course of their disease, although this is largely dependent upon the location of the tumour. No studies have investigated the validity of the NART as an index of premorbid ability in patients with brain tumours. However, it is possible that, when administered to patients with some degree of dysphasia and/or confusion, the task may not have the same validity as it has been shown to have in studies of healthy individuals. Certainly, NART performance has been shown to be compromised in patients with Alzheimer's disease, with NART

performance becoming progressively worse as disease severity increases (O'Carroll et al., 1995). Therefore, since no validation studies have been carried out to date, the NART may not be a feasible tool with which to estimate premorbid ability in patients with brain tumours.

Successful performance on cognitive tests often depends upon a large number of discrete functions. However, many standardised cognitive tests also rely heavily on the intactness of more basic functions such as motor and sensory and abilities, in addition to the participant having a normal level of consciousness (Taphoorn and Klein, 2004). Since a number of brain tumour patients experience mild focal neurological deficits especially of language, motor function and/or vision, successful performance of even short relatively tests may be impeded if the patient does not have suitable vision with which to see the stimulus materials, or has a hemiparesis that prevents successful completion of paper and pencil tasks (Zbinden et al., 2006).

Inaccurate test data may be gathered as a result of emotional distress on the part of the patient and is a further problem associated with carrying out cognitive assessments with neuro-oncological patients. Emotional distress can result in reduced attention, vigilance and motivation, and each of these functions play an integral role in successful completion of almost all cognitive assessments (Taphoorn and Klein, 2004). The issue of heightened emotional distress must be given particular consideration when assessing cognitive function in patients with brain tumours. Kilbride et al. (2007) report heightened levels of anxiety in brain tumour patients after surgery and before commencing radiotherapy as measured by the Hospital Anxiety and Depression Scale (HADS), with a greater prevalence of anxiety specifically reported in younger patients. Furthermore, Grant et al. (1994) found that depression has a negative impact on performance on tests assessing memory and language domains (Williams Delayed Recall Test and Boston Aphasia Severity Rating Scale) and Pringle et al. (1999) report that patients with a meningioma experienced higher levels of anxiety and depression as measured by the HADS than those with any other type of tumour. The study also suggested that levels of both anxiety and depression were found to be significantly lower after surgery compared

to measures taken before surgery, as measured by the HAD Scale. Since the presence and extent of emotional distress is likely to fluctuate throughout the patient's journey, this could have a significant effect on the results of serial cognitive assessments, with poorer performance during periods of heightened anxiety perhaps reflecting greater emotional disturbance as opposed to treatment effects.

Klein et al. (2001) suggest that recruiting a comparison group of patients with a non-CNS cancer diagnosis is a useful means of controlling for the unique stressors associated with a diagnosis of cancer and their potential effects on cognitive test performance. However, in a detailed study, Andrewes et al. (2003) explored the emotional and social function of brain tumour patients in the post-surgical period. Following surgery, patients completed the Emotional and Social Dysfunction questionnaire (ESDQ). The 'partner' version of the questionnaire was completed by the patient's spouse or someone with a similar relationship who knew the patient well. The brain tumour patients and their partners were compared with 49 self-rating controls with a terminal cancer diagnosis who had undergone surgery to extra-cerebral areas and 44 partner-related controls. The brain tumour group comprised patients with astrocytoma (n=13), meningioma (n= 26), neuroma (n=13) and pituitary adenoma (n=17). It was found that in the post-surgical period, brain tumour patients suffered more emotional and social dysfunction as measured by the ESDQ than the matched control group. Specifically, post-hoc analyses revealed significantly increased self-rating of the astrocytoma patient group on the anger, helplessness, inertia, fatigue, indifference, euphoric and inappropriate scales, compared with controls. Moreover, when the brain tumour patients were divided more generally into benign and malignant tumour groups, the malignant group rated themselves significantly worse on the self-report and partner versions of the ESDQ. This study questions the utility of recruiting different groups of cancer patients to serve as a control group (Klein et al., 2001), since the brain tumour patients in this study reported higher levels of emotional dysfunction than controls who had a similar prognosis and surgical treatment. Thus, the potential effect of the unique emotional distress that brain tumour patients experience may not be eliminated by recruiting other cancer patient groups for comparison. However, an adequate control group is

extremely important to allow these effects to be controlled for as much as feasibly possible.

Comprehensive assessment of cognitive function requires a battery of tests that assess a number of different cognitive domains including language, attention, memory and executive functions (Taphoorn and Klein, 2004). However, completing a detailed battery of cognitive assessments is time-consuming and may therefore fatigue or be particularly stressful for patients with brain tumours. In turn, increased fatigue or feelings of stress may reduce motivation and, in turn, negatively influence test scores. Additionally, patients are less likely to take part in studies that place considerable demands upon the participant and as a result, test scores may be unrepresentative (Scotland et al., 2009). Therefore, some studies have employed short cognitive screening tools, such as the Mini Mental State Examination (MMSE) as a measure of cognitive functioning, particularly when assessing the cognitive effects of radiotherapy on brain tumour patients (Brown et al., 2003). However, the shortcomings of the MMSE as a measure of cognitive dysfunction in brain tumour patients have been highlighted (Meyers and Wefel, 2003). The MMSE was initially devised as a brief screening tool for dementia and although it is sensitive to severe cognitive dysfunction, scores in the "normal" range (a score $\geq$ 27/30 points) may not necessarily reflect intact cognition. Moreover, Meyers and Wefel (2003) suggest that stable MMSE scores over time do not necessarily reflect no significant change in cognitive function. The MMSE specifically assesses aphasia, apraxia, orientation and attention. However, a number of other cognitive functions have been shown to be impaired in brain tumour patients, including learning, memory, executive function and processing speed. Therefore, brief assessments of cognition such as the MMSE are likely to be insensitive to cognitive impairment in a large proportion brain tumour patients and there is therefore a need for more sophisticated and discriminating neurocognitive assessment tools for use in cohorts of patients with brain tumours (Klein and Heimans, 2004).

A number of cognitive tests comprise only one version and as such are unsuitable for repeated assessment since practice effects may artificially inflate scores on

subsequent testing sessions. Practice effects have been found to be particularly evident on tests that have a large speed component, require a unique and unfamiliar response mode or have a single solution and practice effects are of particular importance in tests of memory since repeated assessment inevitably results in learning of test materials (Lezak et al., 2004). Some tests have specifically been designed for repeated assessment and have alternate forms that significantly reduce practice effects. However, the proportion of cognitive tests that fulfil suitability criteria for repeated assessment is small since these alternate forms of the same test must have interform reliability (Lezak et al., 2004). Thus, practice effects are an important consideration in many studies that conduct serial cognitive assessments in brain tumour patients, for example, to assess the effect of a specific treatment regime.

Weitzner and Meyers (1997) suggest that, in order for neurocognitive tests to be useful in clinical trials involving brain tumour patients, they must be short, repeatable, sensitive to change and a test battery should measure different brain functions. However, due to the aforementioned problems inherent in the measurement of cognition in this group, many commonly-used methods of assessment fall short of this standard. It would therefore be desirable to have an assessment method that not only avoids the plethora of problems inherent in measuring cognition in brain tumour patients, but also measures an important function.

1.8 Brain Tumours and Information-Processing

In addition to the specific cognitive deficits many brain tumour patients experience throughout the course of their illness, many patients also present with a generalised psychomotor or information-processing slowing, with or without focal neurological deficits. Until recently, no useful assessment tool had been identified as a means of measuring the extent of this slowing in information-processing that is common to patients with brain tumours. However, Zbinden et al. (2006) performed a preliminary study to examine whether inspection time, a measure of the efficiency of the brain's

visual information processing, was impaired in patients with intracranial tumours. Twenty-three newly diagnosed brain tumour patients who were to undergo surgery to biopsy or resect the tumour were recruited into the study and were compared with 24 age and sex-matched control patients who were having elective spinal surgery. Both groups of patients completed a battery of neuropsychological tests that assessed a number of different aspects of cognition, including the Rey Auditory Verbal Learning Test (a measure of learning and memory), National Adult Reading Test (premorbid function), Digit Symbol Coding (processing speed and attention), Trail Making Test Part B (mental flexibility) and the Edinburgh Functional Impairment Tests (EFIT; which provide a measure of memory, speech and limb function). Visual information processing was assessed by the inspection time test. The inspection time test is a two-alterative, forced choice, computer-based measure and the test methodology and background is fully described in Chapters 2 and 3. Patients completed these tests pre-operatively and repeated a number of the measures, including inspection time, in the post-operative period. To compare pre-operative test scores, general linear modelling was used with age included as a covariate since inspection time scores were found to correlate with patient age. Patients with brain tumours had significantly lower inspection time scores than the control patients, indicating a slowing in visual information processing at the time of presentation. However, no significant differences were found between the brain tumour and spinal surgery groups on any of the other cognitive measures that were administered in the pre-operative period. Moreover, the spinal surgery control cohort did not differ significantly in terms of pre- and post-operative inspection time score, with all patients scoring within ± 5% of their pre-operative score when tested at follow-up. However, in the brain tumour cohort, although there was no significant overall change in post-operative inspection time score, 70% of the cohort either increased or decreased their IT score by between 5 and 20%, with 9 patients showing significant improvement and 7 deteriorating significantly. This preliminary investigation thus suggests that visual inspection time may be a potentially useful measure in clinical neuro-oncology. The task itself overcomes many of the problems associated with measuring cognition in neuro-oncological patients and the specific advantages the task holds as a method of assessment in brain tumour patients and other clinical

populations and a review of the inspection time literature are discussed further in Chapter 2.

2 Inspection Time

2.1 History of Inspection Time

The study of individual differences in cognitive ability, and the underlying causes of these differences, is an area of research that has been of interest to both scientists and psychologists for many years. Galton (1883) and Spearman (1904), both pioneers of research into intelligence, believed that individual differences in mental ability could be explained by differences in performance of the most simple of psychological functions (Deary and Stough, 1996). More recently, research in this area has returned to the theoretical stance of these authors, with reductionist efforts used to examine the structure of human cognition and mental ability. This area of research subscribes to the idea that studying the most *basic cognitive processes* will further our understanding of individual differences in intelligence, since these basic processes correlate with higher mental abilities. Reaction time paradigms, and other tasks that involve measuring the speed with which a simple task is performed, have commonly been employed by researchers in an attempt to explain human intelligence differences and a small but consistent association between reaction time and psychometric intelligence has been found (Deary and Stough, 1996). However, there exists some controversy over exactly what are the fundamental processes that reaction time tasks measure and task results can be confounded by motor problems and practice effects (Widaman and Carlson, 1989). These reaction time studies have therefore been unable to identify a basic process that can account for a significant proportion of the variance in human intelligence (Deary and Stough, 1996).

Given the limitations of reaction time paradigms in determining the underpinnings of human cognitive abilities, the inspection time task has attracted a great deal of interest in this area. Of the tasks that claim to measure the most elementary cognitive processes, inspection time has perhaps proven to have the greatest potential to account for the differences in human intelligence. The task itself correlates significantly with a number of different standardised tests of psychometric

intelligence. Moreover, it has a clinical utility and has been used to assess the early stages of visual information processing in a number of different clinical and non-clinical groups. The neuroanatomical correlates of the inspection time task itself have been studied in functional MRI (fMRI) studies and the task has also been administered to groups of children, and employed as a potential tool to explain the observed deterioration in general cognitive function with increasing age in older adults. This chapter will therefore discuss: the relationship between inspection time task performance and higher mental abilities; the neuroanatomical substrates of the task; inspection time studies in children and older adults; the potentially heritable nature of the task; and the utility of the inspection time task as an assessment measure in a number of different clinical groups.

2.2 Method of Testing

It is useful to first give an overview of what the typical method of inspection time testing involves. Measuring an individual's inspection time is a relatively straightforward process and this immediately gives the task an advantage over many other, more complex methods of cognitive assessment. There exist several different methods of testing, however, the two forms of a typical inspection time stimulus are shown in Figure 2.1. Each stimulus shape comprises two parallel, vertical lines joined by a shorter horizontal bar at the top. The stimulus is shown in each trial in the form of one of the versions shown: it has either a longer leg on the left or a longer leg on the right. Discriminating between which line is longer (i.e. left or right) is a simple task for most individuals provided they have adequate visual function. However, if the stimulus is presented for a very short amount of time and is subsequently covered by a backward mask, this can make discriminating between the lengths of the lines very difficult indeed. The backward mask used in the inspection time task consists of a jumble of vertical lines that create a forest-type mask (see Figure 2.2). This mask essentially prevents the stimulus from being processed any further. The time between the beginning of the stimulus presentation and the onset of the backward mask is known as the stimulus onset asynchrony (SOA; Sadler and Deary, 1996).

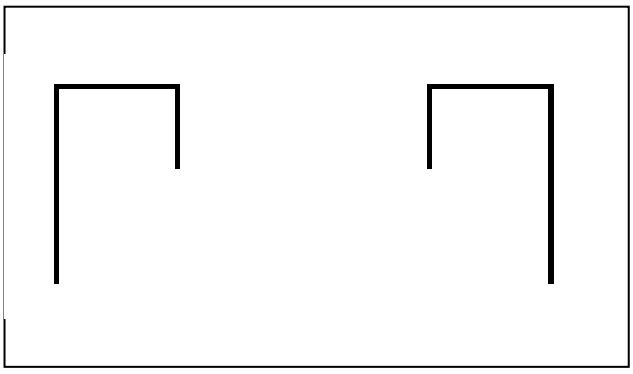


Figure 2.1. The two possible stimulus shapes shown during each inspection time trial.

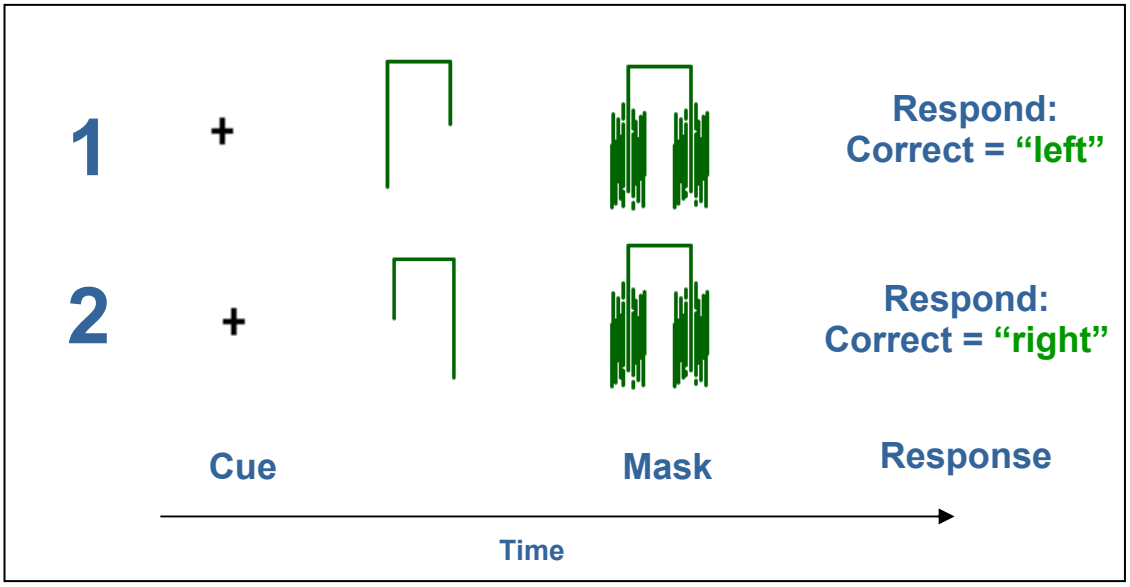


Figure 2.2. Two possible sequences of events in each inspection time trial

In early inspection time research efforts the task was presented tachistoscopically (Vickers et al., 1972, Nettelbeck and Lally, 1976). A tachistoscope is an instrument that can be used to project images onto a screen for precise, varied periods of time, and is accurate to a few milliseconds. In the present day, a computer with a fast vertical refresh rate is used to present the stimuli and to record participants' responses.

The inspection time task comprises a number of trials, the sequence of events for each single trial can be seen in Figure 2.2. In each trial the participant is shown a warning cue (cross) in the centre of a computer screen. This cross is then replaced by one of the two stimulus shapes which are, in turn, covered by a backward forest-type

mask to prevent any further processing of the stimulus. In each trial, the participant is required to indicate, by means of pressing a button on the computer keyboard, which of the lines was longer (left or right). The method of administration of the test varies between studies. The commonly-used version of the task includes a large number of experimental trials, in which the stimuli shapes are shown randomly at differing durations in order to provide an estimate of the participant's inspection time. In the version of the task employed in the present study, the durations for which the shapes are shown ranges from 200ms to 6ms, with 10 presentations at each duration. The participant undertakes a number of practice trials prior to the test trials. This allows a detailed explanation of the task to be given and the participant is given the opportunity to become familiar with the task prior to undertaking the test trials.

Participants are instructed that there are no requirements to respond quickly, and that responses should be made at leisure. A participant's inspection time is often expressed as the duration required for a given level of accuracy in responding to be achieved, e.g. 85% or 90%. However, a total accuracy score of correct responses from a total of 150 is also frequently used to measure inspection time. Usually data from those participants who score <17/20 on the longest two stimulus durations (150ms and 200ms) are considered invalid as most participants should be able to obtain a near perfect score at these durations if the task has been properly understood. Thus, any participant who scores <17/20 on the two longest presentation durations may not have fully understood the task instructions.

The inspection time task is therefore a relatively quick and simple measurement of the early stages of visual information processing that can be applied to a number of different groups of participants.

2.3 Theory of Inspection Time

Although the inspection time task was designed to be essentially theory-neutral, the underlying rationale of the inspection time task is based largely upon two main theories (Vickers and Smith, 1986). The first of these is a perceptual model of

discrimination that assumes visual information is gathered from the environment in small 'quanta' (Vickers, 1970). The second theory is termed the 'accumulator model' of decision making. This model postulates that, in order to make a forced choice discrimination, an individual will take a series of samples of the available sensory information until a critical amount of information has been obtained, such that one of the stimulus alternatives is favoured (Nettelbeck and Lally, 1976, Deary and Stough, 1996). This is referred to as an 'optional stopping process'. Optional stopping models assume that responses will reflect the level of caution with which the decision is made. Thus, this type of model will reflect the speed-accuracy trade-off. Vickers et al. (1972) proposed that, in some instances, a single observation might provide enough information from which to make a forced-choice decision and were interested in how quickly this observation could be made. Thus, the inspection time task was devised and the term 'inspection time' itself was defined as:

"the time required by a subject to make a single observation or inspection of the sensory input on which a decision of relative magnitude is based" (Vickers et al., 1972).

In early inspection time studies, it was initially hypothesised that inspection time would be around 100ms and there would be little inter-individual variation. However, average inspection time was found to be 105ms but a great deal of variation between participants was observed, with estimates ranging from 74 – 144 ms. Since the design of the paradigm is such that the effects of sensory noise and other variables are minimised, it was proposed that inspection time may provide a useful measure of the time taken to make a single observation of sensory input that may be a basic factor that limits perceptual speed and cognitive performance more generally (Vickers and Smith, 1986).

2.4 Inspection Time and Higher Mental Abilities

The idea that individual differences in the speed of mental functioning, as measured by the inspection time task, may be somehow related to individual differences in intelligence, as measured by cognitive and IQ-type tests, has attracted a great deal of

interest. A large number of studies have been carried out to determine whether this theory is supported and inspection time does account for some of the variance in human mental ability. These studies have been carried out using a variety of different clinical and age groups and studies have used a variety of different cognitive and intelligence tests as surrogate measures of intelligence. Moreover, different versions of the inspection time task have been devised, including an auditory version as well as the more frequently employed visual version of the task (Deary et al., 1989, Parker et al., 1999).

Nettelbeck and Lally (1976) were among the first authors to report a significant correlation between inspection time scores and performance on the Wechsler Adult Intelligence Scales. This was an extremely significant finding which paved the road for research in the years to come, and this research continues to the present date. However, given that their study recruited only a small number of participants (n = 10) who had widely ranging IQ scores, the relationship between inspection time and IQ (IT-IQ) was likely to have been somewhat exaggerated. As a result, subsequent studies questioned whether the finding could be replicated within the general population. Nettelbeck and Kirby (1983) found that the IT-IQ correlation was considerably lower when data from participants with exceptionally low IQ scores were excluded from analyses. However, this was not the case when those with high IQs were not included in analysis. Therefore, results of early studies examining the relationship between inspection time and general intelligence were somewhat contradictory. Lubin and Fernandez (1986) concluded that the variability in results across different studies meant that no definitive conclusions could be drawn with regards to the potential existence of the IT-IQ relationship.

Kranzler and Jensen (1989) carried out a meta-analysis, accumulating data from all studies that had examined the IT-IQ relationship until this date. Their primary aim was to determine whether the IT-IQ relationship does in fact exist, and if so, to estimate the size of the relationship between inspection time and intelligence. They included 31 separate studies, both published and unpublished, with 1,120 participants in total. The meta-analysis found that inspection time is (negatively) related to IQ,

and this relationship is particularly strong with measures of general IQ and performance IQ. For adults with general IQ scores, an uncorrected correlation of -.30 with inspection time scores was obtained. After correcting for artifactual error sources (e.g. sampling error, attenuation and range restriction), the correlation was -.54. This meta-analysis therefore provided strong evidence to support the existence of the IT-IQ relationship.

More recently, the above-described meta-analysis has been updated in a further examination of the relationship between inspection time and IQ (Grudnik and Kranzler, 2001). More than 90 published and unpublished studies were included, with a total of 4,200 participants. Again, this meta-analysis found evidence to suggest there is a substantial relationship between inspection time and IQ. These authors report that, in the total sample, an uncorrected correlation of -.30 between inspection time and IQ was obtained. When artifactual effects of sampling error, error of measurement and range variation were corrected for, the correlation between inspection time and intelligence was found to be -.51 and this relationship was generally the same for adults and children (-.51 and -.44 for adults and children, respectively). Grudnik and Kranzler (2001) also extended the previous meta-analysis by addressing two new questions with regard to the IT-IQ relationship. Firstly, they examined whether the relationship is the same across different types of inspection time task - auditory inspection time and the more commonly employed visual version of the task. Auditory versions of the inspection time task have been developed in an attempt to determine whether the IT-IQ relationship is unique to measures of visual information processing or whether it is a general property of the sensory systems. Early versions of the auditory inspection time (AIT) task involved presentation of two differently pitched tones at durations of between 200msec and 6msec. The participant was required to indicate in which order the tones were presented, i.e. high – low or low – high (Deary et al., 1989). More recently, Parker et al. (1999) devised an AIT task that they proposed would overcome the limitations associated with the original version – predominantly its reliance on the participant's ability to discriminate between frequencies which was reported to be excessively difficult for a number of participants (Irwin, 1984). The revised version of the AIT required the

participant to indicate the apparent location (left or right) of a tone that was delivered via headphones to produce a 'dichotic phantom'. The duration of a backward masked tone was varied in order to provide an estimate of auditory inspection time. Grudnik and Kranzler (2001) found the mean corrected correlations for visual and auditory inspection time tasks with IQ measures were -.49 and -.58 respectively and this suggests that the relationship is similar for both versions of the inspection time task. Thus, the IT-IQ relationship has been well-established and it would appear that the relationship exists for both visual and auditory versions of the task.

Attempts to explain the significant correlation between inspection time and intelligence have looked upon inspection time as measuring the speed of a single mechanism such as "sampling input" (Nettelbeck, 2001). Brand is a strong supporter of the theory that inspection time measures inter-individual differences in speed of perceptual processing and it is this basic "mental speed" construct that underpins individual differences in intelligence (Brand, 1996). However, although some authors subscribed to this idea that individual differences in inspection time (i.e. differences in 'mental speed') are in part responsible for individual differences in general intelligence, some researchers take an opposing stance and propose that better scores on inspection time tasks (i.e. quicker speed of processing) is simply the result of having higher intelligence scores (Howe, 1988). A number of theories to support this proposal have been postulated. It has been suggested that the IT-IQ relationship may exist simply because more intelligent individuals are more highly motivated to complete the inspection time task (Mackintosh, 1998), or that those with a higher IQ tend to be less anxious during such tasks, which may have a facilitating effect on inspection time performance (Irwin, 1984). A further proposal is that more clever participants tend to form strategies that facilitate performance on even simple cognitive tasks such as the inspection time paradigm (Grudnik and Kranzler, 2001). These theories would all point to inspection time being simply another task in which clever people perform better. However a number of other studies have generally refuted this idea that good inspection time performance is a result of being more intelligent. The most commonly reported strategy employed by participants is the use of apparent motion cues, described by Luciano et al. (2005) as

a flicker radiating from the bottom of the shorter line on appearance of the mask. Egan (1994) found that although participants who reportedly employed these cues had shorter (i.e. faster) inspection times, they did not also have higher IQ scores. Furthermore, the aforementioned meta-analysis suggests that strategy utilisation in fact impedes inspection time performance, providing evidence against the idea that better inspection time performance is the result of having a higher IQ (Grudnik and Kranzler, 2001). Moreover, further support for the idea that it is inspection time differences that account for some of the variance in IQ scores comes from the use of more effective backward masking procedures in the inspection time task, which render the task more resistant to strategy use. Evans and Nettelbeck (1993) found that the use of backward masking procedures in the IT task has no effect on the obtained correlation between inspection time task score and intelligence. Finally, Stough et al. (1996) refute the idea that inspection time and IQ are related because of personality and/or temperament variables, reporting that personality factors indicative of a more motivational temperament did not mediate the IT-IQ relationship. It would therefore appear that there exists little or no evidence to suggest that strategy use and personality factors play a significant role in the IT-IQ relationship.

Further evidence points to the fact that it is differences in the efficiency of information processing, as measured by the inspection time task that can explain a significant proportion of human intelligence differences. Deary (1995) carried out a study in which children (n = 104) were given an auditory inspection time task and IQ tests at age 11 and again at age 13. Three possible models were tested: that AIT at age 11 causes later IQ at age 13; that IQ at age 11 causes later AIT at age 13; or that there is equal reciprocal causation. It was found that the first model was the best fit and inspection time score at age 11 was found to predict IQ score at age 13 but that the reverse was not true. This finding therefore supports the theory that inspection time accounts for some of the observed individual differences in IQ. There is little evidence to suggest that it is IQ that causes inspection time scores, and this study suggests that the converse is more likely to be true, and inspection time may account for approximately 20% of intelligence test variance (Deary, 1995).

Therefore, a substantial amount of evidence exists to suggest that inspection time accounts for a substantial minority of individual differences in human intelligence and research has now been directed towards examining whether inspection time is differentially related to specific aspects of cognitive function.

2.4.1 Inspection time and factors of intelligence

Carroll (1993) found that there are several group factors within cognitive ability and these include crystallized intelligence, fluid intelligence and processing speed, among others. These group factors load on a higher-order general intelligence ('g') factor (Petrill et al., 2001). Johnson et al. (2008) found evidence to support the idea that there is a unitary general intelligence construct ('g'), in a study that found the correlations between the 'g' factors obtained from five different cognitive test batteries were consistently high. Together with the results of a previous, similar study (Johnson et al., 2004b), this study not only supports the idea of a general intelligence factor but also suggest that most assessments of cognition identify a common underlying component of general intelligence. Since it is now generally accepted that there is a highly significant association between inspection time and measures of psychometric intelligence, research efforts have shifted focus towards specific examination of whether inspection time is differentially related to different types of cognitive abilities and, in particular, whether inspection time is related to intelligence at the general 'g' level or whether it contributes to group factors. The strength and indeed the very existence of the IT-IQ correlation have been shown by numerous studies to be dependent upon the tests used for comparison, with certain cognitive ability measures correlating more strongly with inspection time than others (Burns et al., 1999).

The IT-IQ relationship is stronger for measures of performance IQ ('fluid' intelligence) as opposed to tests that measure verbal IQ ('crystallized' intelligence, (Kranzler and Jensen, 1989). Deary (1993) recruited 87 participants from a diabetic outpatient clinic, each of whom completed the Wechsler Adult Intelligence Scales

(WAIS) and the inspection time task. The participants were otherwise healthy and had varied educational and social backgrounds. Three stages of analysis including correlations, exploratory factor analysis and confirmatory factor analysis revealed that although inspection time had a near zero loading on a verbal factor, it loaded highly on the WAIS subtests that measure performance IQ. Inspection time therefore appears to be significantly related to performance IQ, but is only weakly related to verbal IQ. These results support previous studies that have found inspection time to be significantly more strongly related to measures of performance IQ (Kranzler and Jensen, 1989). Although a clinical population comprised the sample for this study, full-scale IQ scores had an approximate near-normal distribution, suggesting that the sample were likely representative of the general population although future studies should seek to confirm Deary's (1993) findings in a cohort of healthy participants.

Similarly, Osmon and Jackson (2002) examined the relationship between inspection time performance and three specific aspects of intelligence using three factors from the Woodcock-Johnson Battery of Cognitive Ability (Revised) – the visual processing factor (Gv), the fluid intelligence factor (Gf) and also the crystallized intelligence factor (Gc). Forty college students (19 male and 21 female) completed inspection time testing, in addition to six tests from the aforementioned battery in order to obtain the three factor scores. The six tests included Visual Closure and Picture Recognition (visual processing; Gv); Picture Vocabulary and Oral Vocabulary (crystallized intelligence; Gc); and Analysis-Synthesis and Concept Formation (fluid intelligence; Gf). Correlations, partial correlations and multiple regression were used to evaluate any relationships between inspection time and the three intelligence factors studied. There was found to be a moderately strong correlation between inspection time and the visual processing factor and inspection time and the fluid intelligence factor. However, after controlling for its relationship with fluid intelligence, the relationship between inspection time and the visual processing factor was no longer significant. The partial correlation between inspection time and fluid intelligence remained significant after controlling for the relationship with visual processing. To predict inspection time using fluid, visual processing and, crystallized intelligence factors, forward stepwise multiple

regression was carried out. This showed only the fluid intelligence factor to be a significant predictor of inspection time. Therefore, this study offers further support to suggest that inspection time is specifically related to a measure of fluid intelligence (i.e. performance IQ). However, the findings of this study must be accepted with a certain degree of caution given the relatively small sample size with an age range (mean 23.61 years, SD = 7.04) and a likely restricted range of intelligence, with low IQ participants under-represented in the study since the participants comprised college students. The authors themselves suggest that due to a combination of these factors, the large correlation found to exist between fluid intelligence and inspection time may be the result of sampling error. Additionally, the authors acknowledge that the mask used in the inspection time task may have resulted in inspection time values being inaccurately low since it has previously been shown that use of this traditional form of the mask results in around 50% of college students employing an "apparent motion" strategy which makes it possible to perceive quicker presentation rates. Therefore, given the sample of students recruited into the study, the inspection time values obtained cannot be taken as an accurate reflect of inspection time in these participants. Thus, it is proposed that future work should involve replication of the study using an auditory as opposed to visual version of the inspection time paradigm. Should a similar relationship between inspection time and fluid intelligence be found when an auditory version of the task is used, the hypothesis that inspection time reflects some fundamental components of fluid intelligence could be accepted with greater confidence (Osmon and Jackson, 2002)

Petrill et al. (2001) examined whether inspection time is related to psychometric intelligence through group factors, general factors, or both. Five hundred and sixty-eight identical and fraternal twins were recruited, ranging in age from 6 to 13 years. Each participant was tested on three separate occasions. The 11 subtests from the Wechsler Intelligence Scale for Children – Revised (WISC-R) were administered as a measure of psychometric intelligence. Elementary cognitive tasks, including simple-choice reaction time, Stimulus Discrimination (SD), Probe Recall (PR) and Tachistoscopic Threshold (TT), all from the Cognitive Abilities Test battery were also administered to each participant. The Tachistoscopic Threshold task is a

computer-based task in which two diagrams were presented simultaneously for a brief duration and were then masked. Participants were required to determine whether the two diagrams were the same or different and the task was used as a measure of visual information processing, instead of the inspection time task, in this study. After correcting for the effects of age and sex, it was found that inspection time (as measured by the Tachistoscopic Threshold task) and other elementary tasks (SD, PR and simple-choice reaction time) predict general intelligence. In keeping with results of studies described above, it was also found that inspection time predicts performance ability independent of other elementary tasks. Elementary task scores were found to account for 31% of the variance in general psychometric ability ('g') with 26% of this variance related to a factor comprised of motor-based tasks and inspection time. The other 5% of the variance in 'g' accounted for by elementary tasks was found to be related specifically to inspection time. In terms of the variance in performance (fluid) IQ, 3% of the variance is related to inspection time, but not other elementary tasks. Petrill et al. (2001) propose that these findings suggest inspection time may tap variance that is unique to performance (fluid) and general IQ that is not measured by motor-based tasks such as reaction time. Future studies that involve administration of other, more commonly-used versions of the inspection time task, as opposed to relying solely on the Tachistoscopic Threshold task, as well administering a larger battery of intelligence tests, would help to confirm these preliminary results.

Therefore, it has been proposed that inspection time correlates most strongly with measures of fluid intelligence and this has been based upon the premise that performance IQ tests from the WAIS provide a measure fluid ability. However, Burns and Nettelbeck (2003), in a study that aimed to fit the inspection time within a model of cognitive ability, concluded that performance IQ measures from the WAIS do not in fact measure fluid ability. Ninety adult participants were administered a battery of psychometric and chronometric tasks including inspection time, and measures of reaction time. It was found that inspection time loaded on a general speed factor and a general ability factor, but not on a fluid ability factor. Moreover, the study found that none of the performance IQ measures from the WAIS loaded on

fluid ability and the authors therefore conclude that inspection time does not measure fluid ability. The finding that inspection time was related to a general speediness factor and also to a general ability factor suggests that information-processing speed may be common to all cognitive abilities, and this theory should be investigated further in future studies (Burns and Nettelbeck, 2003).

Therefore, in more recent years, there has been a shift from testing the strengths of the IT-IQ relationship with research efforts now focusing on: examination of the processes that are involved in performing the inspection time task; the types of mental ability that correlate most strongly with inspection time; the reasons for the inspection time - ability correlations; the biological bases of inspection time performance; and practical applications of inspection time testing.

2.5 Neuroanatomical Substrates of Inspection Time

Given that inspection time scores have consistently been found to account for a substantial minority of the variance in human intelligence, research has shifted focus to examine the biological basis of the task, with studies attempting to elucidate the specific brain processes that may be involved during visual information processing, as measured by the inspection time task.

2.5.1 Functional MRI Scanning Studies

Further research has since been directed towards obtaining more detailed information regarding the brain processes involved in the inspection time task. Deary et al. (2001) examined whether carrying out the inspection time task is associated with activation of particular brain areas. In a preliminary investigation detailing whether individual differences in inspection time and IQ test performance are associated with activation in the same brain areas, functional MRI scanning (fMRI) was carried out on seven participants during concurrent administration of the inspection time task. During scanning and observation of inspection time stimuli, participants were directed to make their responses mentally, i.e. no physical record of their discriminations was made. A block mode design, using three blocks was employed for the inspection time

testing. One block was referred to as the 'null' condition in which no inspection time stimuli were shown, only the cue and the backward mask were presented. The second block was the 'hard' condition, in which the stimuli were shown for a duration of only 40ms. The third block was an 'easy' condition where the stimuli were shown for 200ms on each occasion. The preliminary data obtained in this study showed a pattern of brain activation and deactivation during completion of the inspection time task in this group of well-educated, healthy participants. 'Activated' areas are described as those brain regions with a higher amplitude blood oxygen-level dependent (BOLD) signal in the 'hard' compared with 'easy' conditions. 'Deactivated' areas were described as brain regions showing a lower amplitude BOLD signal during the 'easy' condition, compared with the 'hard' condition. It was found that, during the "hard" compared to "easy" versions of the IT task, several areas were activated, and several deactivated. Activated areas included the anterior hemispheric fissure region, the precentral cortex bilaterally near the vertex, small amounts in both posteroinferior parietal regions, both anterolateral frontal cortices and both posteroinferior frontal regions. Similarly, several areas were found to be deactivated when responding to 'hard' versus 'easy' inspection time stimuli. These areas include bilaterally in the anterior frontal cortices, extending into the inferior frontal cortex and the left posterior temporal cortex. Deactivation in the medial parasagittal cortex extending up from the medial occipital cortex into the parietal region was also observed, alongside a small deactivated area located in the right and left frontal cortices. These activated and deactivated areas fit the researchers' hypothesis that those brain regions shown to be involved with higher cognitive activities by previous studies (Duncan et al., 2000, Esposito et al., 1999) were similarly affected when performing the inspection time paradigm.

Following these preliminary results, Deary et al. (2004) carried out a second fMRI study of the functional anatomy of the inspection time task in an attempt to overcome some of the significant limitations associated with the pilot study. These limitations included: a small sample size; pre-warning participants whether they would receive easy, hard or control stimuli; limited practice testing on inspection time; and finally the failure to collect participant responses to the inspection time stimuli during

imaging. Therefore, this follow-up study recruited 20 participants, each of whom undertook three inspection time test sessions. The baseline session took part prior to MRI imaging in order to allow familiarisation with the test and the other two inspection time tests (160 trials in each) took place while the participant was in the MRI scanner. The fMRI design for these tests was event-related. The blood oxygenation level-dependent (BOLD) response was computed as a function of the eight levels of inspection time stimulus duration (from 6ms – 150ms) and also as a function of the behavioural responses made by the participant. The aim of the study was ultimately to show that the cerebral processing network underlying inspection time task performance alters depending on task difficulty, as determined by the duration for which the stimulus is shown. Bilateral activation in the inferior fronto-opercular cortex, superior/medial frontal gyrus, and anterior cingulate gyrus was observed during presentation of the quicker, more difficult inspection time stimuli. Additionally, a number of posterior brain areas showed a signal decrease as a function of shorter (i.e. more difficult) stimulus duration and these areas specifically include the left inferior occipital gyrus, the left posterior mid-temporal gyrus and the right posterior inferior temporal gyrus. These patterns of activation and deactivation were consistent with those identified in the pilot study (Deary et al., 2001). Functional connectivity analysis suggested that there exist two separate networks in the brain that are associated with carrying out the inspection time task. The first anterior network comprises the fronto-opercular area, intrasylvian area, medial frontal gyrus and the anterior cingulate cortex. A second, posterior network was also identified and this network includes the precuneus, the posterior cingulate gyrus, the occipital gyrus, and the right superior inferior/mid temporal gyrus comprising sensory-related regions. This posterior network responded more strongly to inspection time stimuli that were shown for longer durations (i.e. ‘easier’ stimuli). These more easily processed stimuli can be considered as “informative percepts”. Some of the aforementioned regions are known to be involved in visual and visuospatial processing. However, others brain areas, including the inferior parietal lobe posterior cingulate, are activated during less specific, fundamental information-processing functions and it would therefore appear that, when performing the

inspection time task, these visual areas are fused with the less specific areas in order to process more informative percepts (Deary et al., 2004).

Conversely, the anterior network identified through functional connectivity comprises brain areas that are activated as task demands increased (i.e. when inspection time stimuli were presented for shorter amounts of time). The authors propose that this network may be fused together in order to process degraded percepts and/or when associative skills are necessary to complete a task. Thus, activation of this anterior network may explain in part the IT-IQ association. This study therefore offers insight into the specific neural networks that are essential to inspection time task performance, in a cohort of young healthy adults.

2.5.2 Neurochemical Correlates

However, despite the important contribution of brain imaging data to inspection time research, fMRI studies do not allow us to infer what the neuronal processes that underpin inspection time performance are. Examining the neurochemical correlates of inspection time is essential in attempts to describe a complete biological model of inspection time.

Early research efforts in this area evolved from studies that found a nicotine-related improvement in inspection time performance (Stough et al., 1995). Thirty-five regular smokers were recruited and asked not to smoke or drink caffeine for 2 hours prior to the testing session. Each participant completed 3 separate testing sessions in which they completed the inspection time task under sham smoking, no-smoking or nicotine smoking conditions. There was found to be a significant effect of smoking status, with significantly improved inspection times obtained under smoking conditions than those in the sham or no smoking conditions. This finding is consistent with the results of other studies in which other measures of processing speed were used (Bates et al., 1994). The authors therefore suggest that nicotine may enhance visual information processing by enhancing stimulus sampling. Subsequently, research was directed towards examination of the potential role of the

cholinergic system in mediating the IT-IQ relationship, given that nicotine acts as an agonist that stimulates the release of the neurotransmitter acetylcholine (Thomson et al., 2000). In their review of work primarily carried out in their own laboratories, Stough et al. (2001) therefore examine the relative contribution of a number of key human neurotransmitters and their effect on inspection time performance in healthy volunteers. In a number of different studies inspection time performance was measured before and after modulating different important central nervous system neurotransmitters and receptor systems. The review concludes that the studies implicate the neurotransmitter acetylcholine and the cholinergic system in successful inspection time performance. Other psychometric tests also seem to rely upon the cholinergic system, although the serotonergic and dopaminergic systems may also be involved in performance of other psychometric measures. These results are further supported by studies that have reported cholinergic system disturbance and impairment in inspection time performance in patients with Alzheimer's disease. It is proposed that future research should focus on examining both the neurochemistry and brain imaging together while the inspection time task is being performed in order to further understanding of the biological basis of the inspection time task.

Studies have also examined the influence of induced hypoglycaemic states on inspection time performance and this has offered further insight into the specific biological processes involved in the task. Given the evidence from controlled studies of experimentally-induced hypoglycaemia that have shown hypoglycaemia to impair the function of the central nervous system, Strachan et al. (2001) carried out a controlled study of the impact of hypoglycaemia on the functioning of the peripheral nervous system. Sixteen healthy individuals were tested on two separate occasions – one under eugylcaemic (i.e. 'normal') conditions and one the other occasion, under induced hypoglycaemia. A practice session was also carried out in order to familiarise participants with the test procedure. Under both conditions, participants completed a cognitive test battery, which included inspection time testing. The motor nerve conduction velocities and the amplitude of the motor action potentials were also measured during each session. Results indicated that hypoglycaemia caused impaired cognitive performance and information processing. Inspection time was

significantly slower during induced hypoglycaemia, as were components of a reaction time test. However, since motor conduction velocity and motor action potentials were not affected by hypoglycaemia, the authors conclude that while multiple levels of information processing functioning may be altered during hypoglycaemic state, peripheral nerve functions are unaltered and it may be the case that the peripheral nervous system does not reflect the central nervous system's need for glucose. However, this study provides evidence against previous propositions that peripheral nerve conduction velocity differences may explain in part the association between psychometric intelligence and information processing ability (Rijsdijk and Boomsma, 1997). Since moderate hypoglycaemia was found to have a negative effect on cognitive test scores, reaction times and inspection time performance but peripheral nerve conduction velocity remained unaffected, this refutes the idea that speed of conduction in the peripheral nervous system reflects the speed of nerve conduction in the central nervous system and peripheral nervous system speed cannot be used as a measure of speed of information processing and intelligence. However, further studies are required to confirm this finding.

Therefore, current research has identified the potential role of different neurotransmitters in mediating inspection time performance. However, future research is needed to gain further insight into the neurochemical underpinnings of inspection time and other cognitive processes in order to develop an accurate neurochemical model of inspection time and other, more complex cognitive functions (Stough et al., 2001).

2.6 Inspection Time in Children

There is now a general consensus that the efficacy of processing speed explains, at least in part, the correlation between inspection time and IQ in adult participants. However, as Anderson et al. (2001) observe, there has been some controversy with regards to exactly what inspection time measures in children. Nettelbeck and Lally (1979) carried out one of the first studies examining inspection time in child participants and they observed a reliable decrease in inspection time (i.e. shorter

exposure time required to make an accurate judgement) in children with increasing age and this finding has since been replicated by Nettelbeck and Wilson (1985) who recruited 10 school children from each of seven different school grades (ages 7 – 13 years). Each child completed inspection time testing and the 7 and 11 year old children repeated the task 2 years later. A different sample also completed the task on two separate occasions, approximately 2 weeks apart. A practice-related improvement was observed in the group of children who were tested twice with a two week-interval between sessions. However, this change was not nearly as significant as the change seen in the groups who were tested over the two year period. It is therefore proposed that the observed longitudinal improvements in inspection time were due to maturation changes, perhaps developmental changes in speed of processing.

The theory that improvement in processing speed with increasing age in children could account for changes in inspection time performance has not, however, been universally accepted. Anderson (1986) and Anderson et al. (2001), in a partial replication of a study by Nettelbeck and Wilson (1985), argued that experience of the inspection time task had a much larger effect on children's task performance than did maturation. A later study detailed inspection time performance in children compared inspection time scores and Peabody Picture Vocabulary Test (PPVT) performance in a group of children aged 6-13 years who were tested in 2001 with scores obtained in 1981 from a group of children aged 6 -13 years who attended the same school (Nettelbeck and Wilson, 2004). It was found that, although there was a significant improvement in general intelligence scores in the cohort tested in 2001 (known as the "Flynn effect" (Flynn, 1999), inspection time scores had remained constant over time. Inspection time scores were still significantly correlated with intelligence, as measured by the PPVT. This study therefore suggests that the Flynn effect (i.e. observed IQ gains in children over a duration of 20 years) is not attributable to improvements in processing speed, as measured by the inspection time task. The study offers strong evidence to support the proposal that inspection time provides a measure of a fundamental mental function that is involved in human intelligence.

Whatever this function may be, it appears to be unaffected by environmental influences that likely cause the observed improvements in intelligence test scores.

The finding that both cohorts showed the expected age effects that have previously been observed in studies examining inspection time performance in children (i.e. shorter inspection times with increasing age) provides evidence *against* Anderson, Reid and Nelson's (2001) proposal that task exposure can account for the observed improvement in inspection time with increasing age in children. Further evidence to refute Anderson et al's (2001) theory comes from Nettelbeck and Vita (1992) who found evidence to suggest that although practice can reduce the aforementioned developmental trends in inspection time, it does not remove them altogether. Thus there does seem to be an age-related improvement in inspection time scores that is present even after controlling for previous inspection time task exposure.

Edmonds et al. (2008) carried out the first twin study of inspection time in pre- and post-adolescent children providing further cross-sectional data regarding the way in which inspection time scores change with age. Pairs of twins and single children aged between 7 and 17 years were tested on inspection time in addition to cognitive and neuropsychological assessments. The Wechsler Intelligence Scales for Children (WISC-III) provided the cognitive measures and the NEPSY, a standardised neuropsychological assessment that measures attention/executive, language, sensorimotor, visuospatial and memory domains. A definite improvement in inspection time with increasing age was reported in this cross-sectional participant sample. To facilitate data analysis, the children were grouped into quintiles based upon their age at the time of participation in the study so as the youngest 20% of the sample comprised the first quintile and the oldest 20% comprised the fifth quintile. One-way ANOVA showed a significant effect of age quintile on IT score. Furthermore, in the whole sample, a positive relationship between inspection time score and participant age was found. These age-related changes could not be explained by potential confounders such as birth characteristics or social background. In keeping with results examining the IT-IQ relationship in adults, this study found inspection time to be significantly correlated with IQ score in this cohort of children

aged between 7 and 17 years and the association remained significant after controlling for age. A raw correlation of 0.58 was found between inspection time score and Full-Scale IQ (FSIQ) in the whole sample and this was significant ($p < 0.01$). After correcting for age, the correlation was 0.26 ($p < 0.01$). Inspection time was reported to correlate specifically with tasks assessing attention/executive function, language, sensorimotor skills, and memory subtest scores. Given that an age-related improvement in inspection time and an inspection time-IQ relationship have now been established, Edmonds et al. (2008) propose that research in this area should now focus on determining exactly what the mechanisms that form the basis of these relationships are. These authors propose that this may be done in three ways: by further examination of the neural correlates of inspection time; by searching for specific genes that may contribute to the shared variance between inspection time and intelligence; or through investigation of any environmental factors that may improve the relationship between IQ and inspection time.

2.7 Inspection Time in Older Adults

Age-related decline in cognitive function has attracted a great deal of interest. Attempting to understand the psychological and biological foundations of cognitive ageing is important in order to facilitate the design of useful interventions and treatments to counteract the detrimental effects of a decrease in mental capacity with age (Waiter et al., 2008).

There is a great deal of variance in terms of the rate at which individuals age cognitively and even in a single individual, different cognitive functions show different rates of change. Spatial abilities, reasoning and problem solving, memory and decision-making (i.e. measures of 'fluid' abilities) have been found to show the greatest decline after the age of 50 compared with 'crystallized' abilities (Nettelbeck and Rabbitt, 1992). A number of theories have been postulated in an attempt to explain the basis of this age-related cognitive decline. Salthouse (1996) supports the theory that speed of information processing may play an important role in understanding the processes involved in cognitive ageing. This theory suggests that

with increasing age there is a decrease in the speed with which the most simple of mental operations can be carried out. This decline in mental speed may, in turn, result in a decline in the ability to perform higher-order cognitive functions, since many are underpinned by speed of processing. Zimprich and Martin (2002) found mental speed changes as assessed by a Number Connection test, and changes in fluid intelligence to be significantly correlated (.53 in a sample of 417 older adults assessed over a period of 4 years). Thus, measuring speed of visual information processing in older adults may be clinically useful by providing an assessment of the extent of cognitive decline. A number of different standardised psychometric tests are designed to assess speed of processing in humans. However, the majority are confounded in elderly people by their reliance on additional high-level, more complex functions (i.e. Digit-symbol coding from the Wechsler Adult Intelligence Scales), or involve speeded reactions which are likely to be slower in more elderly people irrespective of the extent of cognitive decline (e.g. reaction-time tasks). Therefore the inspection time task may have a specific utility in this area given its advantages over more traditional tasks that measure processing speed. As such, it has been used as a method of assessment in older adults in order to examine whether slowed visual information processing is related to more general cognitive decline as measured by standardised measures of cognition.

Nettelbeck & Rabbitt (1992) found that, in a cohort of 104 participants aged between 54 to 85 years, performance on a reaction-time measure, the inspection time task and on the speeded coding task (digit symbol coding) from the Wechsler Adult Intelligence Scales (WAIS) accounted for almost all age-related changes in cognitive performance, as measured by a number of different tasks, including measures of general fluid abilities. These findings support the proposal that an age-related decline in information-processing speed may play an important part in explaining age-related cognitive decline (Salthouse, 1996). However, given that chronological age made a significant contribution to scores on two tests of learning and retention, over and above that attributable to processing speed differences, this suggests that poorer information processing speed with age cannot explain all age-related cognitive changes. The authors suggest that further research should seek to determine whether

changes in memory and learning abilities with increasing age occur independently of the age-related decline in information-processing speed.

The aforementioned study by Nettelbeck and Wilson (2004) has also highlighted the potential for the inspection time task to act as a biological marker for cognitive ageing. Different cognitive abilities decline at different rates and speed of information processing is thought to have a significant influence on the rate at which decline occurs. Although a gradual slowing in information processing speed and therefore a decline in cognitive abilities may occur during early adulthood, these abilities are thought to remain generally stable until the sixth decade of life. However, Der and Deary (2006) report that while simple reaction time shows little slowing until around age 50, choice reaction time slows from early adulthood. The rate of decline differs between individuals and reliable biomarkers are being sought in order to predict changes in function as a result of ageing, with efforts particularly focused upon identifying a marker of accelerated cognitive decline (Stern and Carstensen, 2000). Nettelbeck and Wilson (2004) suggest that, given the inspection time task is a proven measure of the speed of visual information processing, is significantly related to general mental ability and has apparent stability across generations (in contrast to many other standardised tests), this may render the task a potential candidate as a biomarker with which to assess cognitive changes with increasing age.

Bonney et al. (2006) examined inspection time performance in older adults who were diagnosed with mild cognitive impairment (MCI) and were therefore deemed to be at greater risk of developing further cognitive decline or dementia. They found that inspection time is significantly increased (i.e. slower) in participants with MCI by comparison with age, sex and education-matched controls, thus suggesting that slowed information processing is common to patients with MCI. Due to the cross-sectional nature of the study, it was not possible to determine whether inspection time performance can predict those at risk of significant cognitive decline, but the study paved the way for future studies that may examine longitudinally whether changes in inspection time over time have any prognostic value.

Waiter et al. (2008) investigated the potential utility of inspection time as a biomarker in a study that examined whether successful cognitive ageing is associated with retaining those functional brain networks that have previously been shown to be involved with successful information processing (and therefore successful inspection time performance) in healthy young people (Deary et al., 2001, Deary et al., 2004). Members of the Aberdeen Birth Cohort 1936 were recruited at age 70 years and were divided into two groups for the purposes of this study. The Aberdeen Birth Cohort is a group of individuals who had completed the Moray House Test No.12 which assesses verbal, numerical and spatial reasoning as well as other mental abilities at age 11. These individuals were followed up many years later at which time they were administered a battery of cognitive tests at age 64, 66 and 68 years. The members of the cohort were separated into two groups, both of which had similar cognitive ability at age 11, as evidenced by similar scores on the Moray House Test. However, the first group comprised members of the birth cohort who had shown 'successful' cognitive ageing (referred to as the cognitive sustainers) and the second group consisted of members who had 'unsuccessful' cognitive ageing (cognitive decliners). Successful and unsuccessful cognitive ageing was determined by scores on Raven's Standard Progressive Matrices. Both the cognitive sustainers and decliners had their BOLD (blood-oxygen level dependent) activation-deactivation pattern assessed during fMRI scanning while performing the inspection time task. Both groups showed a clear pattern of BOLD activation and deactivation when processing the most difficult inspection time stimuli (i.e. those that were shown for very short durations). When participants made a correct response to the hardest stimuli, activation was seen in the medial, precentral and inferior frontal gyri and deactivation in many more posterior regions was observed. When analysed as two subgroups, the cognitive decliners had far fewer significantly activated regions than did the cognitive sustainers. The anterior cingulate region was significantly more activated in the cognitive sustainers as compared to the decliners.

Functional connectivity maps found there to be more extensively association between brain regions in the cognitive sustainers than in the cognitive decliners. The

authors conclude that when performing the inspection time task, a measure on which performance declines with increasing age, older participants whose cognitive ability (in particular non-verbal reasoning ability) has remained relatively intact with age have a BOLD activation-deactivation pattern than is very similar to that observed in younger humans. Conversely, participants whose non-verbal reasoning ability has significantly declined with age have a far less extensive pattern of associations. Specifically, the activation in the anterior cingulate gyrus is significantly different between the cognitive sustainer and decliner groups and this area was previously identified as being significantly activated when more difficult inspection time stimuli were shown to younger people (Deary et al., 2004). This study therefore concludes that, in older adults who have more intact cognitive functions, similar neural networks underlie information processing speed to those identified in younger individuals, and therefore the task warrants further research as a marker of imminent worsening of cognitive function in old age.

Therefore, the inspection time task is a valuable tool with which to assess cognitive decline with increasing age. Current research efforts are attempting to validate the task as a biomarker that may indicate, and even predict, those individuals who will be cognitive sustainers or decliners in older age.

2.8 Inspection Time and Heritability

The genetics of cognition in general has attracted interest over recent years. Research has attempted to determine whether genetic variants influence individual differences in processing speed and also in working memory since both of these functions have been shown to underpin the ability to perform higher level functions.

The heritability of the inspection time measure has been investigated in twin studies to determine whether the association between inspection time and IQ can be explained by a common genetic factor. Luciano et al. (2001) recruited one hundred and eighty-four pairs of monozygotic twins (i.e. twins who share 100% of their genes) and 206 pairs of dizygotic twins (i.e. twins who share approximately 50% of

their genes). Each participant was tested on inspection time and full-scale IQ was assessed by the Multidimensional Aptitude Battery as part of an ongoing study of the genetics of cognition. This study estimates the heritability of the inspection time measure using the 'twin design' and methods of analysis that reduce the large phenotypic variance in inspection time into genetic and environmental components. The twin design method, in which monozygotic (MZ) and dizygotic (DZ) pairs of twins are compared suggests that if the causes of familial similarity are the additive genes that are transferred from parent to child, the correlation in performance of MZ twins is expected to be twice that of DZ twins. This is because DZ twins share twice as many genes as DZ twins and the theory assumes that the environmental influences are the same in MZ and DZ twins. Results of twin correlations for inspection time showed MZ twins to have greater similarity (r = .34) than DZ twins (r = .21). However, since the MZ twin correlation failed to reach that of the test-retest correlation, this suggests that non-shared environmental influences likely play some role in individual inspection time performance variation. Bivariate analyses of the association between inspection time and IQ suggested that a common genetic factor could account for 36% of the variance in inspection time and 32% of the variance in IQ. No significant common environmental factor was revealed. The genetic correlations between inspection time and the IQ measures were higher than the phenotypic correlation. This therefore suggests that the variation in genes that result in faster inspection time have a strong relationship with the variation in genes that are linked to higher IQs. Thus, evidence from this study suggests that inspection time shares a strong genetic relationship with IQ (Luciano et al., 2001). This finding supports the results of other studies that have investigated other elementary measures of processing speed and their relationship to intelligence.

The twin design has also been used to investigate whether genetic or environmental factors make the greatest contribution to variation in inspection time scores and also to determine whether genetic or environmental factors mediate the IT-IQ relationship (Posthuma et al., 2001). Inspection time and IQ data (assessed by the Dutch version of the WAIS 3R) were collected from 688 family members from 271 extended twin families. Variance components analysis revealed that 46% of the variance in

inspection time could be attributed to genetic influences and the remaining 54% was explained by non-shared environmental influences. A significant phenotypic correlation was found to exist between inspection time and Verbal IQ (0.19) and inspection time and Performance IQ (0.27) and a common genetic factor accounting for 10% of the genetic variance in Verbal IQ and 22% of the genetic variance in Performance IQ was entirely responsible for the correlations.

Thus inspection time appears to be genetically correlated with intelligence (Luciano et al., 2001, Posthuma et al., 2001). Luciano et al. (2004) therefore addressed the question of whether the observed genetic variation in inspection time is associated with intelligence through a unitary factor that influences diverse measures of processing speed and IQ. Monozygotic and dizygotic pairs of twins completed measures of inspection time, choice reaction time and IQ (measured by the verbal and performance subtests of the Multidimensional Aptitude Battery and Digit Symbol Substitution from the Wechsler Adult Intelligence Scale – Revised). It was found that the covariation among processing measures (inspection time and choice reaction time) and IQ test scores could best be explained by a model that comprises a general and specific genetic factor structure, a shared environmental factor that influenced all tests except inspection time, and unique environmental effects that were specific to individual measures. They therefore concluded that a unitary factor is unable to sufficiently account for the relationship between measures of cognitive speed, including inspection time, and standardised measures of cognitive function.

Edmonds et al. (2008) also used the twin design to investigate the genetic contribution to inspection time, on this occasion in children under the age of 16 years. The methodology used in the study is described in section 2.6. Briefly, Edmonds et al., (2008) tested pairs of twins and single children aged between 7 and 17 years old on measures of inspection time, standardised IQ tests (WISC-III) and measures of neuropsychological functioning (NEPSY). In addition to reporting the age-related inspection time changes, genetic analyses of inspection time and IQ were carried out. The heritability of inspection time was estimated to be 45% in childhood and early adolescence suggesting that a high proportion of the variation in inspection

time in this age group is genetic. The best-fit model for the sources of the covariance in inspection time and IQ was one containing both genetic and unique environmental factors and this supports previous findings of the few studies that have examined the heritability of inspection time using the twin-design (Luciano et al., 2001, Posthuma et al., 2001).

Despite the potential limitations of the twin design, namely that mono- and dizygotic twins may not be representative of the general population and thus results may not be applicable to the non-twin population, it would appear that a high proportion of the variance in inspection time is heritable. Research in this area is still in the early stages, and Edmonds et al. (2008) suggest that future work should be directed towards identifying the specific genes that contribute to the IT-IQ relationship.

2.9 Inspection Time in Clinical Populations

The inspection time task has proven utility as a measure of visual information processing that has provided an insight into the processes that underpin variance in intelligence. However, the task is also potentially useful as a clinical tool that can be applied to a number of different populations.

The fundamental process measured by inspection time has been shown to be disrupted in a number of clinical conditions, including Alzheimer's disease, Parkinson's disease, and hypoglycaemia, among others. The inspection time task has a number of advantages over other commonly used measures of processing speed and cognitive ability in clinical populations. The apparent simplicity of the task and the lack of requirement for intact motor or speech function, render the task easier and potentially less stressful for cognitively impaired patient groups to perform than are many other similar standardised tests, at the same time as providing an informative measure that is significantly correlated with cognition.

2.9.1 Inspection time and Dementia

The inspection time task has been used to provide an insight into the extent of slowed information processing and its relationship to cognitive function in a group of participants with mild cognitive impairment in the aforementioned study from Bonney et al.,(2006) The task has also been applied to a group of Alzheimer's disease (AD) patients and a group of patients with Korsakoff's psychosis by Deary et al. (1991) and has been proven to be a useful tool with which to gain insight into the processes underlying AD. The study involved inspection time and other cognitive and psychometric testing of IQ in groups of patients with Alzheimer's disease and Korsakoff's syndrome. The rationale behind the study was that the early stages of visual processing had yet to be thoroughly studied in these groups and inspection time yields a continuous variable that could be useful in identifying disease progression in these groups. The results showed that the group of patients with pre-senile Alzheimer's disease had significant impairment in the early stages of visual information processing (i.e. had poorer inspection time scores) compared to the Korsakoff's syndrome patients and also by comparison with a group of intelligence-matched control participants. This finding is particularly important as impairment in the early stages of information processing had not previously been identified in AD patients. Furthermore, as the authors highlight, it has generally been accepted that the correlation between inspection time and IQ tends to be stronger in groups with poorer cognition. However, although the Korsakoff's patients showed clear impairment on a number of the cognitive tasks, they had very similar inspection time scores when compared to control participants. The authors suggest that this may be because the cognitive deficit in patients with Korsakoff's syndrome arises as a result of impaired processing capacities that are 'downstream' of the information processing function that is measured by the inspection time task.

2.9.2 Inspection time and Parkinson's disease

Patients with Parkinson's disease (PD) have also been the focus of inspection time research. One of the first studies in this area was carried out to investigate the hypothesis that bradyphrenia (slowing of mental abilities) contributes to the

cognitive deficits seen in PD patients (Shipley et al., 2002). The study used the inspection time loop task as a measure of information processing. The inspection time loop task differs from the traditional inspection time task in that it overcomes the potential issue of lack of attention by presenting stimuli in a 'loop', with each repetition of the loop comprising the same stimuli in the same order. The task requires the participant to judge the temporal order of four single letter stimuli. The duration for which the stimuli were presented ranged between 100 and 700ms. Thirty-two patients with PD and 31 age-matched controls were administered the National Adult Reading Test (NART), as a measure of premorbid ability, and completed the inspection time loop task. It was found that the group of Parkinson's disease patients had a significantly lower mean score on the inspection time loop task than did the control group ($p = 0.02$). This poorer ability to discriminate the temporal order of stimuli was apparent even after controlling for pre-morbid IQ, as measured by the NART. These results suggest that there may be a slowing of information processing abilities in PD that is associated with slowed intake of sensory information (Shipley et al., 2002).

Johnson et al. (2004a) utilised the more traditional version of the inspection time task as a measure of elementary cognitive function in patients with Parkinson's disease. Intact motor functioning plays a vital role in successful completion of traditional reaction time paradigms that are often as administered as a measure of "basic" cognition. Since motor function is often disrupted in PD patients, inspection time may provide a suitable alternative measure of elementary cognitive function in this group of patients. A group of "optimally medicated" PD patients were compared with a group of healthy age-matched controls on the inspection time task. In the first part of their study, results indicated that PD patients required significantly longer stimulus presentation time in order to make a correct judgment than did the healthy control group, suggesting that visual information processing is significantly slowed in this patient group. In a second experiment, a group of PD patients were tested on inspection time both when they had and had not taken their usual dopaminergic medication and were compared with age-matched controls. A significant impairment on inspection time was observed in the PD group, irrespective of whether they had

taken their medication or not, by comparison with the healthy control group. This study therefore confirms previous findings that suggest there is a significant deficit in perceptual processing in PD patients. Given that the medicated and non-medicated groups had similar inspection time scores, the authors suggest that cognitive slowing in PD patients may not be improved by dopaminergic drugs.

2.9.3 Inspection time and Depression

The inspection time task has also been employed as a measure of processing speed in patients with clinical depression. Tsourtos et al. (2002) used the inspection time task to determine whether speed of visual information processing is slowed in young patients with unipolar depression. Following an earlier study which concluded that younger depressed patients do not show the cognitive slowing that has been observed in middle-aged and older depressed patients (Purcell et al., 1997), Tsourtos et al. (2002) wished to determine whether this finding was an artefact of the cognitive speed measures employed since they relied upon on a number of other cognitive functions, in addition to processing speed. The inspection time task therefore offered a suitable measure of the early stages of visual information processing that is unaffected by motor speed or other cognitive strategies, and may therefore provide a more reliable measure of cognitive slowing in depressed patients. Tsourtos et al. (2002) recruited 20 unmedicated inpatients with a diagnosis of unipolar depression, 19 medicated depressed inpatients and 20 age, sex and verbal-IQ matched healthy controls. Each group completed inspection time testing and inspection time scores were found to be significantly different between the three experimental groups. There was a significant difference between the control group and the unmedicated depressed group but no significant differences between the control group and the medicated depressed group were found in post-hoc analyses. The unmedicated group had significantly slower inspection times than the medicated group. Therefore, in contrast to previous findings this study showed that speed of information processing, as measured by the inspection time task, is impaired in young, unmedicated, unipolar depressed patients. That the medicated, depressed patients had significantly faster inspection times than the group of unmedicated depressed patients suggests that the

cognitive slowing that is associated with a diagnosis of depression may be somewhat ameliorated by medication. The authors therefore propose that assessment of cognitive slowing should be included in the neuropsychological profile of depression in patients of all ages in future studies of depression. The study also emphasises that the potential effects of medication should be considered in studies that detail cognitive function in depressed patients.

2.9.4 Inspection time and Schizophrenia

Visual backward masking (VBM) tasks that employ a methodology similar to that of the inspection time task have frequently been employed as a measure of early information-processing deficits that have consistently been observed in individuals with schizophrenia (Braff et al., 1991). Butler et al. (1996) tested both medicated and unmedicated schizophrenic patients on a VBM task and found that no significant differences in between the performance of the two groups, despite the fact that symptoms were reported to change on and off medication.

In addition to providing an assessment of the effects of prescribed medications on visual information processing, VBM tasks have also been used to examine the relationship between visual pathway function and symptomatology in patients with schizophrenia. Butler et al. (2002) tested the hypothesis that the VBM deficit in schizophrenic participants is the result of an overactive transient visual pathway response to the mask. Thirty-five patients with a diagnosis of either schizophrenia or schizoaffective disorder and thirty-five control participants completed a traditional backward-masking task in addition to tasks in which the mask had been altered to bias stimulation to transient (low spatial frequency) or sustained (high spatial frequency) channels. Results confirmed a significant deficit on the traditional backward-masking task in schizophrenic patients compared with controls, supporting the findings of the aforementioned studies. The group of schizophrenic patients were also significantly impaired on both the low and high spatial frequency versions of the task by comparison with the healthy control group. There was no difference between

the performance in the medicated and unmedicated conditions, lending further support to Butler et al.'s (1996) findings.

Thus, although no studies to date have used the inspection time task itself as a measure of information processing in schizophrenic patients, similar visual backward-masking tasks have a proven utility as a measure of slowed information-processing in this group. Further research should perhaps consider the use of inspection time in order to gain further insight into the processes that underlie the disruption of visual information processing in people with schizophrenia.

2.9.5 Inspection time and Anaesthesia

The inspection time task has been used as a measure of the effect of nitrous oxide that is used for anaesthesia on psychomotor function in a single study (Cheam et al., 1995). A battery of psychomotor and information processing test measures were completed (inspection time, tapping frequency, critical flicker fusion, picture memory and time sense) by seven healthy volunteers who had been administered various concentrations of nitrous oxide. It was found that, in addition to dose-related memory impairment and reduction in tapping frequency, an increase in inspection time was also observed. That is, visual information processing was slowed as a result of nitrous oxide. No other studies have been carried out to date using inspection time as a measure of processing slowing following a general anaesthetic to confirm the findings of this small study.

2.9.6 Inspection time and Dyslexia

Only a single study has been carried out specifically investigating inspection time performance in dyslexic individuals, although a lot of work has focused on assessing information processing speed in dyslexic individuals, using other measures. Whyte et al. (1985) recruited a group of children who were either dyslexic or normal readers. Each child completed the inspection time task. The dyslexic children were found to have significantly longer inspection times (i.e. slower information processing) than

the group of normal readers. Additionally, there was greater inter-individual variation in the dyslexic group and this group also benefited from more practice trails than did the control group. It is therefore proposed that the inspection time task may measure a process that plays a role in the development of reading skills and that speed of information intake, as opposed to more general perceptual impairment may be involved in dyslexia. Furthermore, it would appear that dyslexic children suffer more task anxiety, given the observed improvement in inspection time performance with practice. It is therefore proposed that the difficulties experienced by dyslexic individuals may result from difficulties at a simple level and more practice opportunities may improve performance on inspection time and other information processing tasks. Given the small sample size and the interesting results of this study, it is surprising to note that no further studies assessing the utility of the inspection time as a measure of visual information processing in dyslexic individuals have been carried out to date.

2.10 Inspection Time – Conclusions

Since its conception in 1972, a great deal of research has been directed towards the inspection time task. The IT-IQ relationship has been established and studies have consistently found that inspection time is strongly related to measures of performance IQ. Studies examining the potentially heritable nature of this relationship are on-going and the neurological correlates of the task are similarly being studied in an attempt to underpin the reasons for which differences in human intelligence exist. Event-related MRI studies are currently being used in older adults, in an attempt to identify the key to successful cognitive ageing. This in turn, may help to identify potential interventions to ameliorate the symptoms of age-related cognitive decline. Moreover, the inspection time task has proven utility as a measure of processing ability in various clinical populations, including patients with Parkinson's disease, Alzheimer's disease and schizophrenia and is increasingly being employed as a clinical tool. This is because the inspection time task itself holds a number of advantages over other standardised tests of cognition: it can be applied across a wide age range, from young children, to older adults; it provides a measure

of a fundamental process and is therefore unaffected by many potential confounding variables, such as motor and/or speech problems. Furthermore, unlike many neurobehavioural tests, it has very limited practice effects since there is no 'content' to the stimuli presented and is therefore useful for repeated administration .The task also offers a relatively quick and easy measure of general cognition which is less stressful and tiring than comprehensive cognitive test batteries that are generally very time-consuming. Thus, the inspection time task continues to attract a great deal of research attention and is becoming an increasingly important measure of information processing ability in a number of clinical populations.

2.11 Aims of Thesis

Given the advantages the inspection time test holds over many other standardised measures of cognition, its potential utility in neuro-oncology is of interest. Most patients with brain tumours are initially treated with surgery (biopsy or resection). Given the limited survival time associated with many primary and metastatic brain tumours it is important that quality of life is not negatively affected by surgical intervention. Most studies of neurosurgical complications associated with brain tumour surgery report physical as opposed to cognitive impairments, despite the fact that cognitive dysfunction can have the greatest impact on quality of life. Furthermore, few studies have assessed cognitive function in brain tumour patients at the time of presentation which makes elucidation of the role of the tumour itself and the effects of different treatments on cognition problematic. Therefore, a measure of visual information processing before and after surgical intervention in neuro-oncological patients would be informative. Following the successful pilot study that demonstrated slowed inspection times in brain tumour patients (Zbinden et al., 2006) this thesis now aims to carry out a larger, prospective study to further evaluate the role of inspection time as a measure of visual information processing ability in patients with brain tumours.

Specifically, the following research questions will be addressed:

(i) Is visual information processing slower in a larger cohort of brain tumour patients prior to surgical intervention, by comparison with age and sex-matched surgical and healthy control groups?

(ii) What are the effects of tumour type, location and laterality on inspection time and other functions?

(iii) Prospectively, what are the effects of surgery on brain tumour patients' inspection time functions alone, and in relation to other neuropsychological measures?

(iv) How do inspection time scores relate to a measure of Quality of Life?

3 Methodology

3.1 Participants

Three groups of participants were recruited into the study. The study group (brain tumour group, n = 118) comprised newly-presenting patients who were admitted to the Department of Clinical Neurosciences (DCN), Western General Hospital, Edinburgh. This group of patients were to undergo a first surgery to biopsy or resect a supratentorial intracranial tumour (either primary or secondary) that had been diagnosed on CT or MRI scanning. The second group (spinal surgery controls, n = 85) were a group of patients who were admitted to DCN for elective surgery to treat degenerative lumbar or cervical spine disease. The third group (healthy controls, n = 80) consisted of healthy volunteer participants who were not receiving any treatment as a hospital inpatient. All three groups were matched as far as possible in terms of age, sex and National Adult Reading Test score (NART, as a measure of premorbid intelligence).

3.2 Patient recruitment

Detailed recruitment procedures were implemented in order to maximise the number of patients recruited into the study. Regular contact was made with relevant medical staff on the DCN wards, including pre-admissions nurses, nurse practitioners and doctors who were responsible for brain tumour and spinal surgery patient care. Recruitment posters detailing the study and relevant contact details for the researcher were placed visibly on the wards to remind staff that recruitment was on-going (see appendix A). The admissions diary for the aforementioned wards was consulted on a daily basis in order to obtain names of potentially eligible elective spinal surgery and brain tumour cases.

The researcher approached any suitable patients at an appropriate time, gave verbal information about the study and offered the patient an information sheet (appendix B). Those patients from the brain tumour and spinal surgery groups who agreed to participate completed the first testing session the day before their operation wherever possible, although in some cases it was only possible to test the patient on the morning of their operation. In a number of other cases, spinal surgery and brain tumour patients were recruited and tested during their visit to the Pre-Admission Clinic (PAC), or Planned Investigation Unit (PIU), up to a week before the date of their operation. This alternative recruitment protocol arose following a change in admission procedures that resulted in the majority of elective surgery patients being admitted on the day of their planned surgery. Thus, approaching and testing patients during their pre-operative outpatient assessment overcame the difficulties associated with recruiting patients pre-operatively.

Brain tumour patients were excluded if:

- They had previously undergone surgery to remove a brain tumour (i.e. not undergoing first craniotomy),
- They had a past medical history that included any psychiatric illness or brain-related disease,
- They did not speak English as a first language,
- They had a severe visual defect (although a number of participants in the brain tumour group had some form of visual field defect),
- They had undergone previous radiotherapy to the brain,
- They were unable to give informed consent (e.g. due to confusion and/or dysphasia).

3.3 Healthy control group recruitment

Members of the healthy control group were recruited from a number of sources, namely the University of Edinburgh's Psychology Department Recruitment Panel. The recruitment panel comprises volunteer members of the public who are willing to be contacted as potential control participants in research studies taking place within

the University of Edinburgh. Permission to use the recruitment panel was sought and granted. Where possible, potential participants were contacted and given information about the study, including the relevant participant information sheet (see appendix C) via electronic mail. Those members of the panel for whom no e-mail address was available were contacted by letter (see appendix D), and were asked to contact the researcher should they be willing to participate in the study. The healthy control group also comprised volunteers recruited from a number of other sources, including volunteer colleagues, friends and family of the research team. Potential participants who had any prior knowledge of any of the tests used were excluded. Healthy control group participants were tested at a mutually convenient time and their travel to and from the hospital was reimbursed, where applicable.

Demographic data including age, occupation, educational background and handedness were recorded for all participants. Data concerning tumour type (WHO classification), location, clinical features, surgery type (biopsy or resection) and medications were also recorded for the brain tumour cohort.

3.4 Procedure

The brain tumour and spinal surgery groups completed a detailed battery of neuropsychological, mood and functional tests prior to surgery (baseline/session 1). Informed consent was given prior to commencing the testing session (see appendix E and appendix F). Completion of the battery in its entirety required approximately 1 hour. Both surgical patient groups were tested on a second occasion, post-surgery and prior to discharge from hospital (approximately 4 or 5 days later). This second testing session lasted about 20-30 minutes. At session two, patients repeated a number of the tests carried out during session 1 (see Table 3.1 for schedule of tests). Members of the brain tumour cohort were tested on a third occasion, where possible. This third testing session usually took place when the patient attended hospital for a follow-up neuro-oncology out-patient appointment, between 10 and 14 days post-surgery. The third session lasted 20 - 30 minutes. The majority of patients who completed the third follow-up session were glioma patients since patients with other

tumour types did not routinely return for follow-up in this initial post-operative period. Additionally, most patients who completed session 3 testing completed only inspection time testing since there was often limited time available for testing prior to the patient's scheduled appointment. Given that patients were given their histological diagnosis at this appointment, it was not feasible to test after their appointment. The healthy control group completed the baseline session at a time of their choosing, and session two was completed between 2 and 10 days later to mirror the pre and post-operative testing of the surgical patient groups. All participants were given the tasks in the same order (see appendix G), with the exception of cases where some tasks were omitted due to patient fatigue, time constraints or patient refusal to complete one or more of the tasks.

Testing sessions were carried out at a variety of different times throughout the day in order to fit in with the demands of the hospital ward and to minimise disruption to the patient. All testing sessions were carried out in the same private, quiet room within the hospital, under consistently artificially lit conditions.

	Tests	Baseline	Session 2	Session 3
Brain Tumour Patients (n = 118)	IT + DS + EFIT	X	X	X
	Other Cognitive	X		
	QOL	X		X
Spinal Surgery Controls (n = 85)	IT + DS + EFIT	X	X	
	Other Cognitive	X		
	QOL	X		
Healthy Controls (n = 80)	IT + DS + EFIT	X	X	
	Other Cognitive	X		
	QOL			

Table 3.1. Schedule of tests. IT = Inspection Time; DS = Digit Symbol-Coding; QOL = Quality of Life Assessment; EFIT = Edinburgh Functional Impairment Tests; Other Cognitive = National Adult Reading Test, Rey Auditory Verbal Learning Test, Trail Making Test Part B, Verbal Fluency, Letter-Number Sequencing, Hospital Anxiety and Depression Scale, Barthel Disability Index, Karnofsky Performance Scale.

3.5 Measures

The following measures comprised the baseline test battery and a number of the tasks were also included in follow-up assessment (session 2 and session 3; see Table 3.1 for details).

3.5.1 Cognitive Assessments

3.5.1.1 Inspection Time

The test was run and responses analysed using E-Prime software. All stimuli were presented on an iiyama computer monitor, running at a vertical refresh rate of

approximately 120 Hz. Participants were seated comfortably directly in front of the monitor, with their eyes approximately 75cm from the computer screen.

The inspection time task comprises a two-alternative forced choice procedure in which participants are asked to make a simple visual discrimination. They must decide which of two parallel vertical lines of visibly different length is longer. In each trial, a small cross is presented in the centre of the computer screen for a duration of 500 ms, serving as a cue. The cue is then replaced by one of the two potential stimulus shapes, which is then immediately covered by a mask (see Figure.3.1). The method of constant stimuli was used, meaning that a fixed number of trials were presented randomly at a number of pre-determined stimulus exposure durations. Therefore, participants completed ten trials at each of 15 stimulus presentation durations, presented randomly, and the task comprised a total of 150 trials. The durations for which the stimulus shape was shown were (in milliseconds): 6, 12, 19, 25, 31, 37, 44, 50, 62, 75, 87, 100, 150 and 200. The 150 trials were divided into 5 blocks of 30 trials, with a break in between each block. This break allowed participants to rest their eyes for a short time, should they so wish. After each trial was complete (i.e. after the mask had been shown), participants were prompted to indicate which line (left or right) they thought was longer (i.e. to indicate which stimulus shape was shown) by pressing one of two keys on the computer keyboard. Each participant was instructed to respond at leisure since response times were not recorded. A number of practice trials alongside standardised, detailed instructions from the researcher were given to each participant prior to presentation of the150 test trials to ensure comprehension of the task and familiarisation with the relevant response keys. A small number of patients in the brain tumour group who had some form of motor weakness or who were anxious at the prospect of using a computer preferred to voice their responses. For this small proportion of patients, the researcher pressed the appropriate button for each response and did so in a location from which it was not possible to see the computer screen, in order to avoid any potential bias.

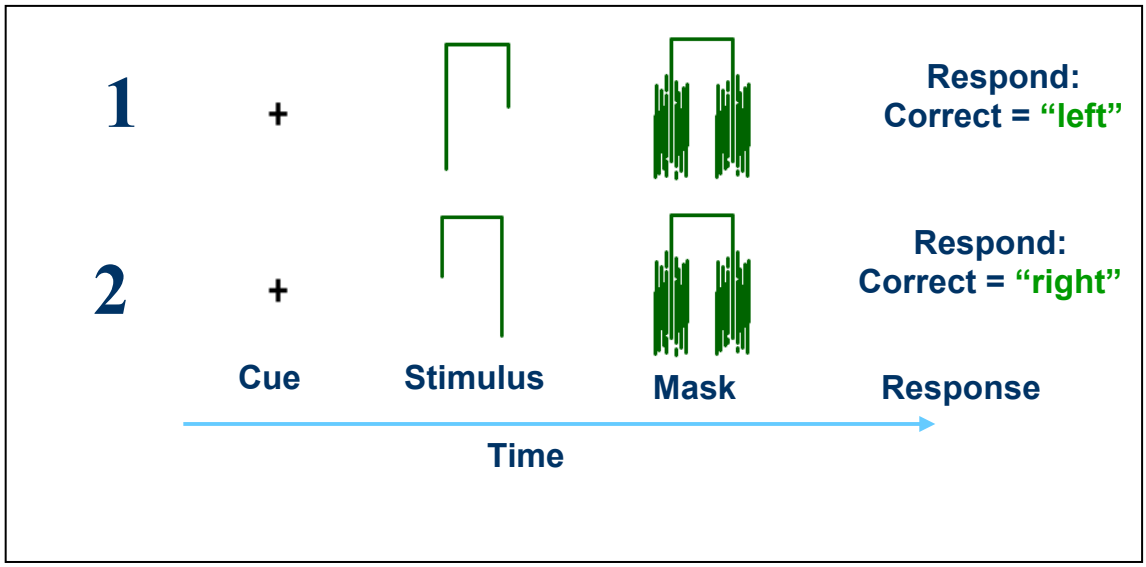


Figure 3.1. Sequence of events in each inspection time trial

3.5.1.2 Rey Auditory-Verbal Learning Test (RAVLT; Rey, 1964)

The RAVLT is a memory test that consists of two lists of 15 unrelated words (List A and List B; see appendix H). List A is read aloud at a rate of one word per second and the participant is instructed to recall as many words as they can, in any order. This procedure is then repeated four times. Thus, there are 5 presentations of List A, with recall of words from the list after each presentation. The 5 List A trials are then followed by a single presentation of the second list of words (List B), with recall of as many words from List B as possible. The participant is then instructed to recall as many words as possible from List A, on this occasion without the list being read aloud. A delayed recall trial of List A is given 20-30 minutes later (following completion of the inspection time task). At this time, the participant is again instructed to recall as many words as possible from List A, again without the list being read aloud. On each occasion, the examiner records each word recalled, in the order the participant recalls them.

The RAVLT provides a measure of learning and retention and is an assessment of many aspects of memory including word-span, acquisition, interference and delayed recall. Word list learning has also been found to be the most sensitive measure of

verbal memory, because there is no context in which to associate the different items to be remembered which increases task difficulty (Lezak et al., 2004).

3.5.1.3 National Adult Reading Test (NART; Nelson, 1982)

The NART is a list of 50 phonetically irregular words, i.e. words that do not follow the usual grapheme-phoneme pronunciation rules, listed in order of increasing difficulty (see appendix I). Participants are required to read the list aloud and are instructed to simply guess how to pronounce any words they do not recognise. Correct pronunciation of the words suggests that the patient has prior knowledge of them and the test was scored as the number of words incorrectly pronounced from a total of 50 on this occasion.

There is strong evidence to support the hypothesis that performance on the National Adult Reading Test (NART) significantly reflects scores on standardised measures of intelligence. Specifically, NART performance at age 77 has been found to correlate significantly with IQ scores obtained from the same participants at age 11. This suggests that the NART may provide a reliable index of prior intellectual functioning (Crawford et al., 2001). Moreover, test scores are relatively unaffected by neurological or psychiatric disorder and as such can provide an estimate of prior, as opposed to current, level of intellectual functioning (McGurn et al., 2004). Therefore, the NART was administered as a measure of premorbid intelligence.

3.5.1.4 Trail Making Test Part B (Reitan, 1958)

This is the second of a two-part pen and paper test devised by Reitan (1958). The test is presented on a single piece of A4 paper comprising a number of randomly placed encircled numbers and letters and the participant is instructed to connect the encircled numbers and letters consecutively with straight lines, alternating between the two sequences (e.g.1-A-2-B-3-C etc, see appendix J). A shorter practice page is completed prior to completion of the actual test page to ensure the participant has fully understood the task instructions. Participants are encouraged to connect the

circles as quickly but as accurately as possible and the score is the time taken, in seconds, to connect all the circles correctly.

The trail making test part B provides a measure of scanning and visuomotor tracking, divided attention and cognitive flexibility although motor speed has a strong contribution to successful, quick completion of this task (Lezak et al., 2004). Poor performance on the test has been found to reflect cognitive inflexibility to a modest degree with test scores correlating more highly with a test of perseveration (Wisconsin Card Sorting Test) than with digit span, letter fluency, or memory scores (Korte et al., 2002). Moreover, both parts A and B of the trail making test have been found to be a sensitive measure of cognitive impairment in people with dementia, even in the early stages of the disease (Botwinick et al., 1988). Performance on both version of the test has also been found to be slower than control participants in patients with mild traumatic brain injuries (Leininger et al., 1990).

3.5.1.5 Verbal Fluency (Controlled Oral Word Association Test; (Benton and Hamsher, 1989)

The Controlled Oral Word Association Test (COWAT) is a measure of verbal fluency that requires the patient to spontaneously produce as many words as possible beginning with a specific letter of the alphabet within a specified time frame (60 seconds). The letters C, F and L were used in the present study. Participants were instructed to verbalise as many words as they could think of beginning with each letter, excluding proper nouns, numbers and repetitions of the same word with a different suffix. A practice trial using the letter 'S' was undertaken prior to the main test and the researcher recorded each word produced during each trial (see appendix K). The score on the COWAT is the total number of words generated for each of the three letters.

The COWAT has been validated with a number of different letter combinations, including 'C, F, and L' and 'F, A, and S' and performance across these different forms of the tasks does not differ significantly in healthy groups, psychiatric patients

and patients with suspected CNS impairment (Troyer, 2000, Lacy et al., 1996). Moreover, written word fluency and phonemic fluency also correlate significantly (Cohen and Stanczak, 2000). Effective performance on this task is thought to involve executive functions that include cognitive flexibility, strategy utilisation, suppression of interference and response inhibition (Abwender et al., 2001). The verbal fluency task is sensitive to brain dysfunction and patients with frontal lesions have been found to generate fewer words on the test, with left frontal lesions resulting in greater impairment than right frontal lesions (Miceli et al., 1981).

3.5.1.6 Digit Symbol-Coding (Wechsler, 1997)

This is a subtest from the Wechsler Adult Intelligent Scales (WAIS-III) and is an assessment of psychomotor performance that is largely unaffected by memory, learning and intellectual ability (Erber et al., 1981). The numbers 1-9 are each paired with a different symbol, shown in a key displayed at the top of the page. Rows of randomly presented numbers with empty boxes below in which the corresponding symbol should be drawn are shown on the same page, below the key (see appendix L). The participant is instructed to draw the relevant symbol under the corresponding number and to complete as many as possible in a 120 second time-frame. The score is the number of boxes filled in correctly within the time limit.

Test-retest reliability is high for this WAIS-III subtest, with correlation coefficients between .82 and .88 obtained in cohorts of healthy participants (Matarazzo and Herman, 1984, Wechsler, 1981). However, test-retest reliability tends to vary when clinical populations are studied, with poor reliability for schizophrenic patients, but coefficients near normal levels for patients with mild traumatic brain injury (Lezak et al., 2004).

3.5.1.7 Letter-Number Sequencing (Wechsler, 1981)

Letter-number sequencing is a subtest of the Wechsler Adult Intelligence Scales (WAIS-III). The patient is read a mixed random combination of numbers and letters

and is required to repeat the numbers first, in ascending order, followed by the letters in alphabetical order. The task increases in difficulty from two items to eight items, with three trials at each level of difficulty (see appendix M). The span is increased until the participant fails all three trials of one length (up to eight characters). Crowe (2000) has shown that performance on this task is related not only to digit span and attentional measures, but is also a measure of processing speed and visual-spatial working memory.

3.5.2 Mood Assessment

3.5.2.1 Hospital Anxiety and Depression Scale (HADS, Zigmond & Snaith, 1983)

This is a short questionnaire which was completed independently by the majority of participants. The questionnaire is a self-screening one for anxiety and depression and was designed specifically to detect emotional disorder in hospitalised patients (Zigmond and Snaith, 1983). Participants are instructed to answer each item based on their feelings over the past week and are also encouraged to give an immediate response and not to spend too long thinking about their answers (see appendix N). The HADS has been found to be a sensitive measure of post-stroke depression although it is reported to be a better measure in men than in women (Aben et al., 2002).

3.5.3 Functional Assessments: Edinburgh Functional Impairment Tests (EFIT)

The EFITs consist of four separate assessments of patient function designed to provide a brief, easily administered assessment of upper and lower limb function, language and short-term memory (Grant et al., 1994).

3.5.3.1 Nine Hole Peg Test

The Nine Hole Peg Test (NHPT) assesses manual dexterity and upper limb function. The participant is required to place nine wooden dowels (9mm diameter, 32mm long) into a wooden base with nine holes spaced 50mm apart as quickly as possible, using one hand at a time and picking up only one peg at a time. A stop-watch was used to record the time taken for each hand separately, measuring the time taken from the first touch of the first peg until the last peg was placed. In cases where the tests could not be completed for one or both hands due to severe weakness or numbness, the test was terminated and a score of 180 seconds recorded. A NHPT score of >18 seconds is considered abnormal.

3.5.3.2 Timed Ten Metre Walk

The Timed Ten Metre Walk (TMW) is a quick and objective test of lower limb function. Participants are instructed to walk as fast as they can, without running, along a straight line ten metre course. The time taken was recorded using a stop-watch and a TMW time of ≥ 8 seconds is considered abnormal.

3.5.3.3 Williams Delayed Recall Test

The Williams Delayed Recall Test (WDRT) is a short assessment of delayed recall (memory) where the participant is shown a piece of A4 paper with nine black and white pictures of unrelated, everyday objects on it. The participant is instructed to have a good look at the pictures and to try their best to remember them. He/she verbally identifies the pictures to the researcher and is then given a few seconds longer to examine the pictures. About 6 minutes later (following completion of the NHPT and Boston Aphasia Severity Rating Scale subtests), the participant is instructed to recall as many of the pictures as they can, in any order. If the patient is unable to recall all nine objects, he or she is given a pre-determined prompt for the missing objects. If any objects are not correctly recalled following the prompt, the participant is then shown 15 pictures, 9 of which are the original ones and is asked to identify the original objects. Scoring is as follows: 2 points for every item not

recalled spontaneously, 3 points for each item not recalled with a prompt and 4 points for each item not recognised from the visual prompt. A WDRT score of >16 is considered abnormal (Clyde et al., 1998). The WDRT has three versions (A, B and C, see appendix O), allowing for repeated administration and a Latin-square method was used to randomise the versions given to each participant at the different testing sessions.

3.5.3.4 Boston Aphasia Severity Rating Scale

The Boston Aphasia Severity Rating scale (BASRS) assesses language and allows the researcher to rate presence and severity of dysphasia on an ordinal scale. Scores on the BASRS range from 6 (normal speech) to 0 (no useful speech or auditory comprehension). Language function was scored on this scale for each patient after listening to conversational speech during the testing session and through formally asking the patient to describe what was going on in a the 'cookie theft' from the Boston Diagnostic Aphasia Examination, see appendix P). A BASRS score of < 6 indicates the presence of some degree of speech impairment and, as such, is considered abnormal.

3.5.4 Quality of Life Assessment

3.5.4.1 European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) + Brain Cancer Module (QLQ-BN20)

The EORTC QLQ-C30 is commonly used to measure health-related quality of life in cancer patients. It comprises five functional scales (physical, role cognitive, emotional and social); three symptom scales (fatigue, pain and nausea and vomiting), a global health status/QOL scale and a number of single items assessing additional symptoms that are frequently reported by cancer patients (dyspnoea, loss of appetite, insomnia, constipation and diarrhoea and perceived financial impact of the disease). The brain cancer module has been developed primarily for brain tumour patients who are undergoing radiotherapy and/or chemotherapy and includes 20 questions that

assess the following: future uncertainty, visual disorder, motor dysfunction, communication deficit, other disease symptoms such as headaches and seizures and treatment toxicities including hair loss (see appendix Q). The majority of participants completed the questionnaire independently, although some brain tumour patients did require assistance to indicate their responses due to either motor problems or mild confusion.

3.5.5 Clinical Assessments

3.5.5.1 Barthel Disability Index (BDI, (Mahoney and Barthel, 1965)

This assessment was completed by the researcher as a measure of physical disability by rating the patients' ability to perform a number of activities of daily living (see appendix R). Ratings were made based on the researchers' own observations, through conversation with the patients themselves and consultation of the patient's notes, in addition to consultation with medical staff involved with the patient's care, where necessary. The Barthel Disability Index yields a score /20, with a score of less than 20 considered abnormal.

3.5.5.2 Karnofsky Performance Scale (KPS, (Karnofsky and Burchenal, 1943)

The KPS is a scale that is often used to assess terminally ill patients and allows classification in terms of functional impairment and medical care requirements. Ratings were made for each patient on the basis of the researchers' judgements and scored on a scale from 100 (no impairment) to 0 (dead, see appendix S). KPS score is often used by clinicians when assessing a patient's suitability for treatment.

4 Recruitment Bias

4.1 Overview

Before reporting the results of the study, it is important to consider the sample of brain tumour patients who were successfully recruited into the study in the context of the entire newly-presenting cohort of patients that were admitted to the department during the recruitment period. Therefore, a recruitment log was kept from the outset of the study to record all newly presenting brain tumour patients who were not entered into the study, and the reason for which they did not participate. This was done to determine whether the sample of brain tumour patients recruited into the study was representative of the total number of new patients admitted to the neurosurgical ward during the course of the study recruitment phase.

4.2 Results

According to the recruitment log, a total of 294 newly diagnosed patients, who potentially fulfilled eligibility criteria for the study, were admitted to the department for a surgical biopsy or resection of a supratentorial brain tumour between April 2006 and January 2009. Of these, 220 (74.8%) were eligible to take part in the study (see Figure 4.1). The remaining 74 patients (25.2%) were excluded from the study for either medical reasons such as severe visual acuity or visual field defects, severe dysphasia or impaired mental status that precluded obtaining informed consent from the patient; or sensorimotor impairment that prevented the patient from being transported to the testing office; or as a result of co-morbidities including Down Syndrome, mental health problems or chronic alcoholism ('ineligible').

Of the cohort eligible to take part in the study (n = 220), only 53.6 % (n = 118) were recruited, since 24.6% of the eligible patients (n = 54) declined to take part. The remaining 21.8% (n = 48) were not entered into the study because they were 'missed' by the researcher. This was due to either annual leave, failure of ward staff to inform

that a new patient had been admitted, or because the patient was admitted or transferred to the ward very late on the day before their operation.

In the total cohort of newly presenting patients who were admitted between April 2006 and January 2009 (n = 294), two patients were recruited into the study and did not subsequently undergo any surgical intervention. These two cases are therefore excluded from the following analysis. In the revised total cohort (n = 292), the majority of patients were subsequently found to have a high-grade glioma (54.8%), low-grade glioma (10.6%), meningioma (16.1%) or metastasis (10.6%) (see Table 4.1). However, in the study cohort (i.e. the group of patients who participated in the study, excluding the two cases who did not have surgery), only 44% had a high-grade glioma and 19.8% had a low-grade glioma. This difference arose because the majority of patients in the medically ineligible group had a high-grade glioma (64.9%), whereas only 4.1% of ineligible patients had a low-grade glioma.

The group of patients who declined, were deemed medically ineligible, or were missed, were combined to form a single 'non-participation' group. The group of patients who did participate were compared with the 'non-participation' group in terms of tumour type using a chi-square test. The analysis revealed a significant association between group and tumour type, $\chi^2(4) = 20.76$, $p < 0.001$. This confirms that fewer high-grade glioma patients than expected were entered into the study, whereas a higher number of low-grade glioma patients than expected did take part (see Table 4.1).

The 'declined' and 'medically ineligible' groups were then combined to create a second 'non-participation' group. By not including the group of 'missed' patients, this makes the comparison more relevant to other, similar studies of cognition in newly-presenting brain tumour patients. Comparison of the participation group against the combined declined and ineligible groups again revealed a significant association between the group and tumour type, $\chi^2(4) = 16.67$, $p = 0.002$. This again demonstrates that a higher than expected number of high-grade glioma patients did

not participate in the study. Conversely, fewer low-grade glioma patients than expected did not take part.

Finally, the participation group were compared with the group of patients who were approached but declined to take part in the study. In this instance, the association between group and tumour type failed to reach statistical significance, $\chi^2(4) = 8.49$, $p = 0.075$.

A paper detailing the potential for bias in neuro-oncological studies, including recruitment data from April 2006 – December 2007 has been accepted for publication (Scotland et al., 2009; see appendix T).

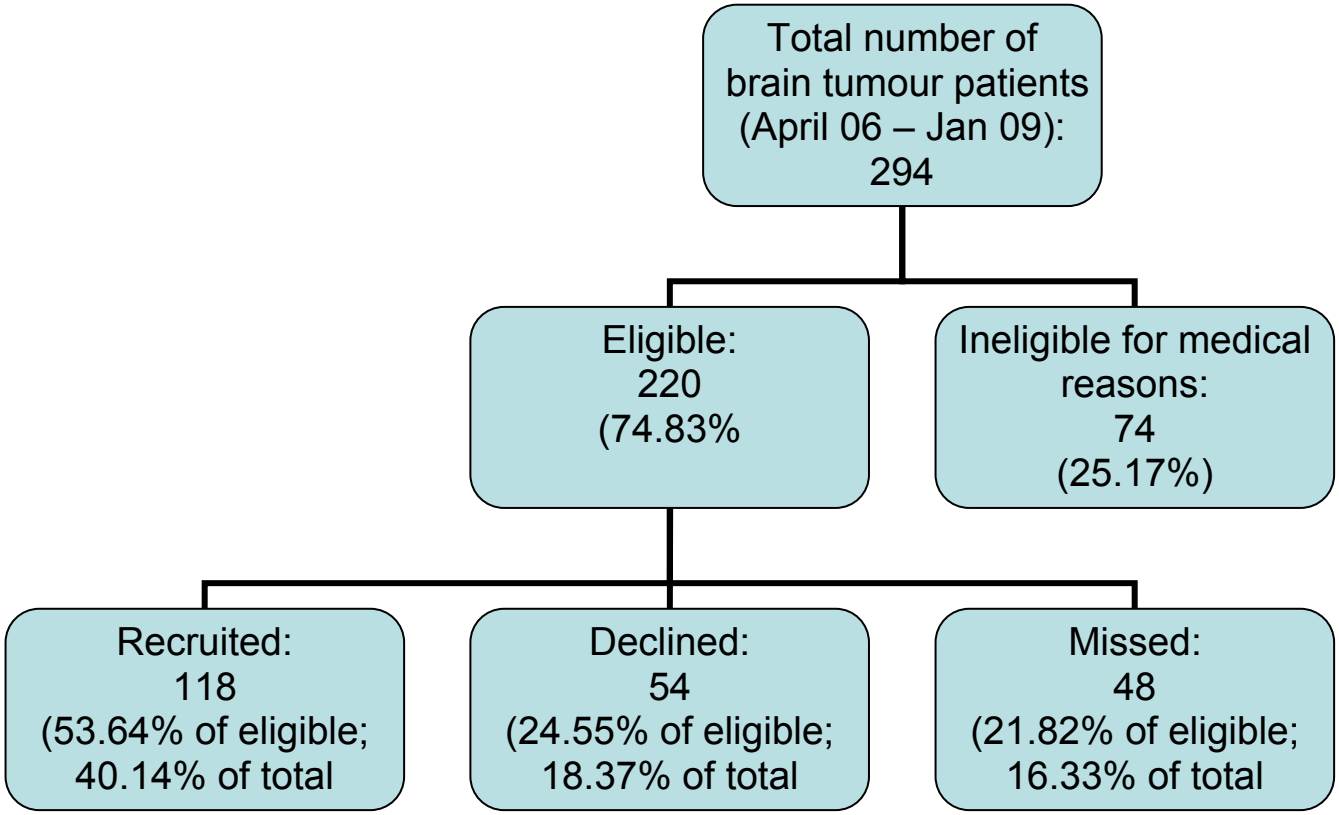


Figure 4.1. Flowchart detailing brain tumour cohort eligibility for the study and recruitment.

Table 4.1. The number of patients with each tumour type in the total cohort (excluding 2 patients who were recruited but did not undergo surgery to obtain a histological diagnosis), and subgroups of those who participated, declined, were ineligible, or missed.

		Tumour Type				
		High-Grade Glioma (WHO III – IV)	**Low-Grade Glioma (WHO I – II)**	**Meningioma**	**Metastasis**	**Other/ Not Known**
Group	**Participated** n (% within group)	51 (44.0%)	23 (19.8%)	17 (14.7%)	14 (12.1%)	11 (9.5%)
	Declined n (% within group)	28 (51.9%)	2 (3.7%)	11 (20.4%)	9 (16.7%)	4 (7.4%)
	Ineligible n (% within group)	48 (64.9%)	3 (4.1%)	11 (14.9%)	6 (8.1%)	6 (8.1%)
	Missed n (% within group)	33 (68.8%)	3 (6.3%)	8 (16.7%)	2 (4.2%)	2 (4.2%)
	Total n (% within group)	160 (54.8%)	31 (10.6%)	47 (16.1%)	31 (10.6%)	23 (7.9%)

4.3 Discussion

Detailed examination of the recruitment log kept throughout the duration of the study revealed that a large number of patients were ineligible for medical reasons (74/294) and a large number of patients who were approached subsequently declined to take part in the study (54/220). Furthermore, a significant number of potentially eligible patients were not approached at all (48/220). There were significant differences between the histological characteristics of the brain tumour patients who were successfully recruited into the study and those who did not participate for any of the aforementioned reasons. By comparison with the combined cohort of 'non-participating' patients who were either ineligible, declined to participate or were not approached ('missed'), the recruited cohort comprised significantly fewer patients with high-grade gliomas than expected and, conversely, significantly more patients with low-grade gliomas than expected did participate in the study. This finding was replicated when only those patients who declined and those who were ineligible were included in the 'non-participation' group. Comparison of the patients who were approached but declined participation with those who did take part in the study did not reach the conventional level of statistical significance.

Recording and examining the characteristics of not only those patients who were successfully entered into the study, but of all potentially eligible patients and detailing the reasons for non-participation is of particular importance in studies such as the present one. Flick (1988) highlights the issue of 'preinclusion attrition' which occurs when potential participants do not participate in the study as a result of some selection bias. Thus, considering preinclusion attrition allows us determine whether specific groups of patients are more likely to participate than others and is an important consideration when considering the representative nature of a patient sample and subsequent results.

The unexpectedly high refusal (decline) rate in the brain tumour group in the present study was in contrast with several other studies that have been conducted in the same unit, including two non-interventional, non-testing but nevertheless demanding

studies (Bastin et al., 2006, Schwindack et al., 2005) and an ongoing non-interventional clinicopathological study (Whittle et al., 2007). However, the cognitive testing demands of the present study may have specifically contributed to the high refusal rate. Certainly, the specific problems associated with neuropsychological testing in brain tumour patients have been documented qualitatively, although few studies report recruitment rates quantitatively (Taphoorn and Klein, 2004). In the present study, it was clear that the anxiety associated with a sudden diagnosis of a brain tumour and trepidation at the prospect of completing neuropsychological tests appeared to be the primary reason for patient refusal, although this was not documented quantitatively. The specific need to recruit patients in the pre-operative period was also particularly problematic and this may explain why relatively few studies have documented cognitive function in brain tumour patients at the time of presentation and prior to any surgical treatment, yet several similar studies have successfully examined cognition following surgery and/or radiotherapy in this patient group.(Bosma et al., 2007, Ek et al., 2005, Klein et al., 2001). However, although the decline rate in the present study was considerably higher than that observed by other neuro-oncological studies that were carried out in the same unit, the successful recruitment rate of 54% of eligible patients (40% of the total cohort) is comparable with studies in other areas, including a community-based clinical research study in which 47% of all eligible participants were successfully recruited (Wiemann et al., 2005). The recruitment rate in the present study was also higher than the mean enrolment proportion (32.7%) across 112 recruitment sites into a Cardiac Arrhythmia Suppression Trial (CAST; (Shea et al., 1992) and higher than the 22.4% of cases successfully recruited into a randomised interventional ISAT study (ISAT, 2002).

A large proportion of newly presenting patients were subsequently found to be ineligible (25%) and were therefore excluded and not approached. This ineligibility was usually a result of acute confusion and/or dysphasia that precluded informed consent. Memory problems, confusion and speech impairment are commonly reported symptoms in neuro-oncological patients and in an audit of Scottish brain tumour patients, Grant (2004) reports that memory problems/confusion were present

in 31% of patients at the time of hospital presentation and dysphasia was noted in 21% of patients. These data are in keeping with the proportion of patients deemed to be ineligible for these reasons in the present study.

Of the total cohort of newly presenting patients with supratentorial brain tumours for whom surgical intervention was planned, 55% had a high-grade glioma, 11% had a low-grade glioma, 16% had a meningioma, 11% a metastasis and the remaining 8% had other tumour types. These figures are broadly in keeping with those reported in the literature (McKinney, 2004). However, the majority of studies reporting the incidence of tumours examine primary brain tumours or gliomas specifically and this makes it difficult to determine the representative nature of the total brain tumour cohort who represented to the neurosurgical wards. However, given the large geographical area covered by the unit, it can reasonably be assumed that the total recorded cohort were representative of the brain tumour population in the UK.

The finding that the group of brain tumour patients who participated have significantly different histological characteristics than those who declined, were ineligible and missed combined, and also by comparison with the combined group of patients who declined or were ineligible patients is an important consideration when extrapolating the results of the present study. Low-grade glioma patients were over-represented and high-grade glioma patients were under-represented in the recruited cohort and this finding is perhaps unsurprising for a number of reasons. Patients with more aggressive tumours (i.e. high-grade gliomas) tend to present with more severe focal and global symptoms that may impair their ability to give informed consent to participate in cognitive assessments. Moreover, patients with suspected high-grade gliomas more often present on an emergency basis and are often admitted to the hospital ward for only a short time before surgery, in contrast with their low-grade counterparts. This procedure can be problematic when the study protocol necessitates patient recruitment and testing prior to surgical intervention. Patients with a suspected low-grade glioma are more likely to be offered surgery on an elective basis, often after a period of ‘watch and wait’ observation (Whittle, 2004). Therefore, given that patients with a low-grade glioma are more likely to have had time to adjust

to the diagnosis of brain tumour, that ample time is generally available to recruit these patients prior to surgery, and these patients tend to have less severe deficits, the number of patients who refuse, are missed or deemed ineligible is likely to be fewer and this may explain the above findings.

The disparity between the proportion of high and low-grade glioma patients in the recruited and non-recruited groups may well bias the final outcome of the present study. Since patients with high-grade gliomas would likely have performed less well than their low-grade counterparts (Bosma et al., 2007), the recruited brain tumour cohort may have performed better than a truly representative cohort and this must be considered when interpreting the results detailed in the following chapters. The potential for consent or participation bias should also be considered in other similar neuropsychological studies in neuro-oncology and these findings have been discussed in our recent paper (Scotland et al., 2009).

5 Baseline/Pre-Operative Function

5.1 Descriptive Data

5.1.1 Demographic Information

Between April 2006 and January 2009, 118 brain tumour patients (60 male, 58 female) were recruited into the study. A further 85 spinal surgery controls (39 male, 46 female) and 80 healthy volunteer controls (31 male, 49 female) were also evaluated during this period. The demographic characteristics of each group are shown in Table 5.1.

Chi-square tests showed that there were no significant differences between the three groups with respect to sex (χ^2 (2) = 2.810, p = 0.245). The mean age of participants in the brain tumour group was 50.0 years (SD 13.0), was 47.8 years (SD 11.2) in the spinal control group, and was 49.1 years (SD 16.4) in the healthy control group. One way ANOVA revealed no significant differences between participant age in the three groups (F(2,282) = 0.681 p = 0.507). Similarly, chi-square analysis showed there to be no significant differences in handedness between the three groups (χ^2 (4) = 4.715, p = 0.318).

Crosstabulation revealed that many more participants in the healthy control group had university degree level qualifications than expected (see Table 5.2), suggesting that the healthy control group had achieved higher levels of education that the brain tumour and spinal surgery groups. Formal testing, using Pearson chi square analysis of the highest levels of qualification achieved by participants in each of the three groups, revealed significant differences between the groups (χ^2 (12) = 103.260, p < 0.001).

The mean National Adult Reading Test (NART) score, included as a measure of premorbid ability, was 20.6 (SD 9.4) for the brain tumour group. The corresponding mean score for the spinal surgery control group was 19.9 (SD 9.3) and for the healthy

control group was 11.7 (SD 5.0). Higher scores on the NART reflect poorer premorbid ability.

As can be seen from Table 5.3, the group of brain tumour patients had a variety of histological diagnoses, with tumours located in a variety of different brain regions. Thirty patients had a biopsy of the tumour, and 86 had a resection. The remaining two patients who were tested at baseline did not subsequently go on to have surgery. All brain tumour patients were taking pre-operative dexamethasone.

5.1.2 Attrition and Follow-Up

Figure 5.1. shows attrition and follow-up rates in the brain tumour, spinal surgery and healthy control groups. The majority of healthy control participants completed both testing sessions (76/80, 95.0%). A smaller number of spinal surgery control patients completed both baseline and post-surgical session 2 testing sessions (67/85, 78.8%). Of the eighteen patients who did not complete the post-operative testing, 10 were approached but refused to participate post-operatively. This was usually the result of increased pain following surgery. The remaining 8 were classified as 'missed' since they were not approached post-operatively, prior to discharge. This was due to the patient having been discharged earlier than expected, without the researcher's knowledge.

In the brain tumour cohort, 64 of 118 (54.2%) participants completed the first post-operative follow-up testing session, prior to discharge from the hospital ward. Those patients who did not complete session 2 either declined to take part at this stage, had acute physical and/or cognitive impairment post-operatively that prevented completion of follow-up testing or were discharged unexpectedly without the researcher being made aware. Of the 118 patients who completed baseline testing, 52 (44.1%) completed the third testing session, which took place when the patient returned for out-patient follow-up. Those patients who did not complete this third session, as with those who were not tested at session 2, most frequently declined to take part. Furthermore, patients with a diagnosis of a meningioma or metastasis were

not routinely offered an out-patient appointment in the department and therefore could not be asked to take part in the third testing session. Therefore, the majority of patients who completed the third testing session had a diagnosis of either high or low-grade glioma. Overall, 80/118 (67.8%) participants in the brain tumour cohort completed at least one post-operative follow-up session. Of the 38 patients (32.2%) who completed only baseline testing, the majority declined to take part in any post-operative follow-up (n = 17, 44.7%), or were unable due to post-operative complications including increased confusion, dysphasia or hemiplegia (n = 12, 31.6%). The remaining patients who took part in only baseline testing were either missed by the researcher due to annual leave (n = 5, 13.2%) or did not have any surgical intervention (n = 2, 5.3%). Comparisons of the baseline performance of those participants who completed session 2 follow-up testing and those who did not are detailed in appendix Chapter U.

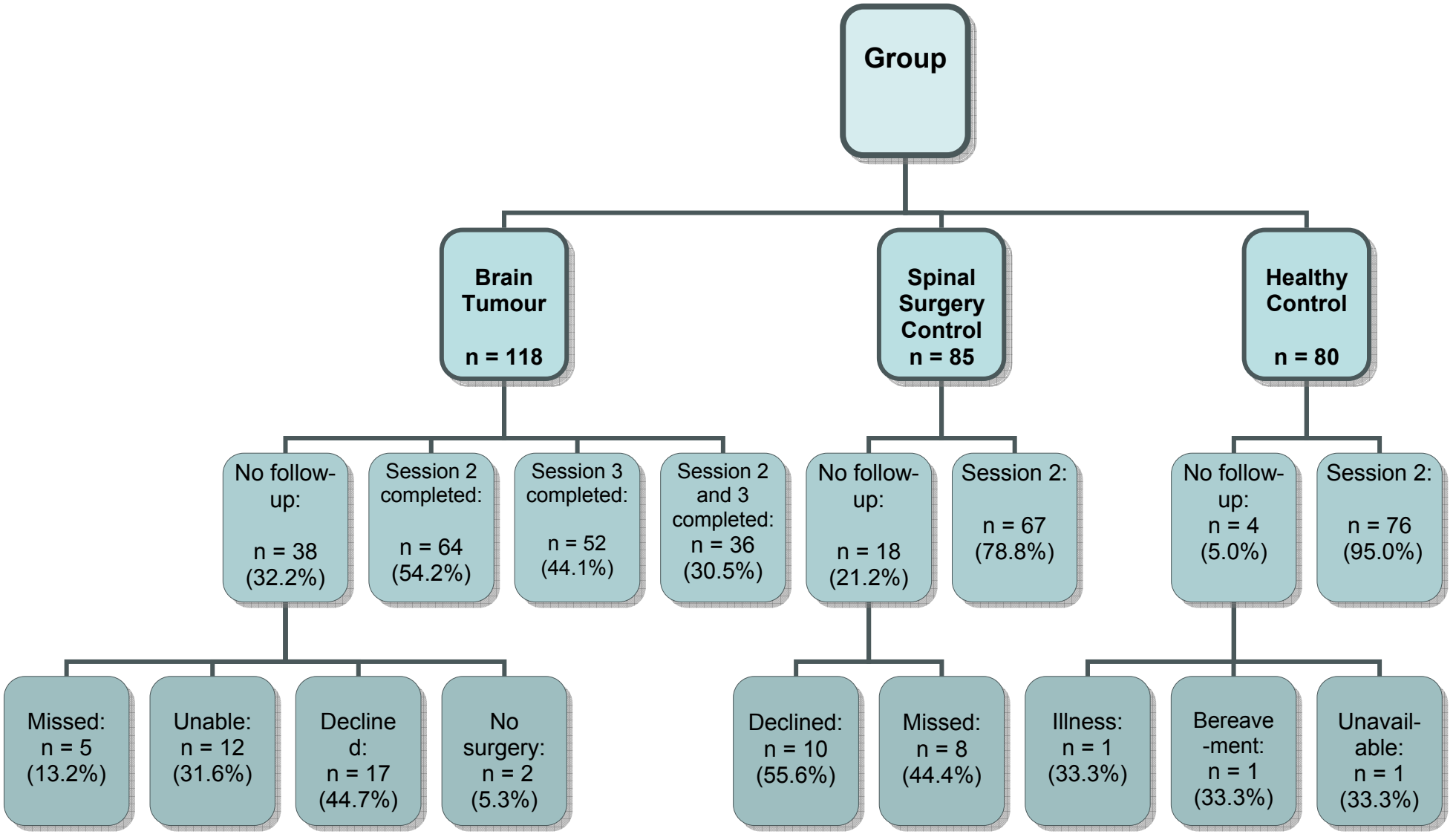


Figure 5.1 Attrition and follow-up rates in the brain tumour, spinal surgery and healthy control groups.

5.2 Baseline Analyses: Cognitive Measures

5.2.1 Overview of Analysis Procedure

Baseline (i.e. pre-operative/session 1) scores on each of the cognitive tests were compared between the brain tumour, spinal surgery and healthy volunteer groups using general linear modelling (analysis of covariance). Group and sex were entered as fixed factors in the models.

In the whole sample (i.e. all three participant groups combined) who met validity criterion in inspection time testing (i.e. at least 17/20 correct on the two longest durations, as explained in chapter 2.2), age was correlated with inspection time ($r (n = 255) = -0.444$; $p < 0.001$). There was also a significant correlation between National Adult Reading Test (NART; as a measure of premorbid ability) score and total inspection time score in the whole sample combined ($r (n = 252) = -0.124$, $p = 0.048$). Therefore, age at the time of testing and NART score were included as covariates in all group comparisons.

In each of the sections that follow, mean scores for each group (brain tumour, spinal surgery control and healthy control) are presented with standard deviations (SD) in brackets. The results of the general linear modelling analysis are then reported. Estimated marginal means (adjusted for age and NART score) and pairwise comparisons are then described for each of the tests administered to highlight the differences between the groups where applicable. Least Significant Difference (LSD) adjustments were used to compare main effects. There was no significant group by sex interaction in any of the models, therefore these data are not reported below.

Table 5.4 shows the results of the general linear modelling analyses for each test and Table 5.5 gives the estimated marginal mean (standard error) scores, adjusted for age and NART score for the three groups, alongside the p-values for significance for pairwise comparisons of the brain tumour group against each of the two control groups, for each test.

5.2.2 Inspection Time Scores: All Inspection Time Data

A total of 114 brain tumour patients, 82 spinal control patients and 80 healthy volunteer controls completed the inspection time task in its entirety at baseline and all inspection time data, including scores from those participants who were deemed to have 'invalid' data (i.e. those who scored < 17/20 correct on the longest two task durations) were included in the analysis in the first instance. Mean total inspection time total scores, with standard deviations shown in brackets, were 114.3 (SD 19.9) for the brain tumour group, 121.4 (SD 13.7) for spinal controls and 126.4 (SD 11.5) for the healthy control group.

General linear modelling, with total inspection time score (i.e. score /150) as the dependent variable revealed a significant main effect of the covariates age, $F(1,266) = 76.95$, $p < 0.001$, partial $\eta^2 = 0.224$; and National Adult Reading Test (NART) score, $F(1,266) = 18.584$, $p < 0.001$, partial $\eta^2 = 0.065$. Older participants and participants with higher (poorer) NART scores were significantly more likely to have lower inspection time total scores. There was no significant main effect of sex in the model, $F(1,266) = 2.604$, $p = 0.108$, partial $\eta^2 = 0.010$. There was a significant main effect of group (i.e. brain tumour, spinal or healthy control) in the model that included the effects of age and NART score, $F(2,266) = 7.994$, $p < 0.001$, partial $\eta^2 = 0.057$.

Pairwise comparisons using LSD tests revealed a significant difference between the brain tumour group and the spinal group ($p = 0.005$) and a significant difference between the brain tumour group and the healthy volunteer group ($p < 0.001$). Estimated marginal mean scores, adjusted for age and NART score, were 115.9 (SE 1.3) for the brain tumour group and 121.6 (SE 1.6) for the spinal surgery group, showing that the brain tumour group had significantly poorer scores than the surgical control group. The corresponding mean score for the healthy control group was 124.2 (SE 1.7). Therefore, the brain tumour group were also significantly impaired on inspection time at baseline by comparison with the healthy control group. There was no significant difference between the performance of the spinal surgery control group and the healthy control group ($p = 0.280$).

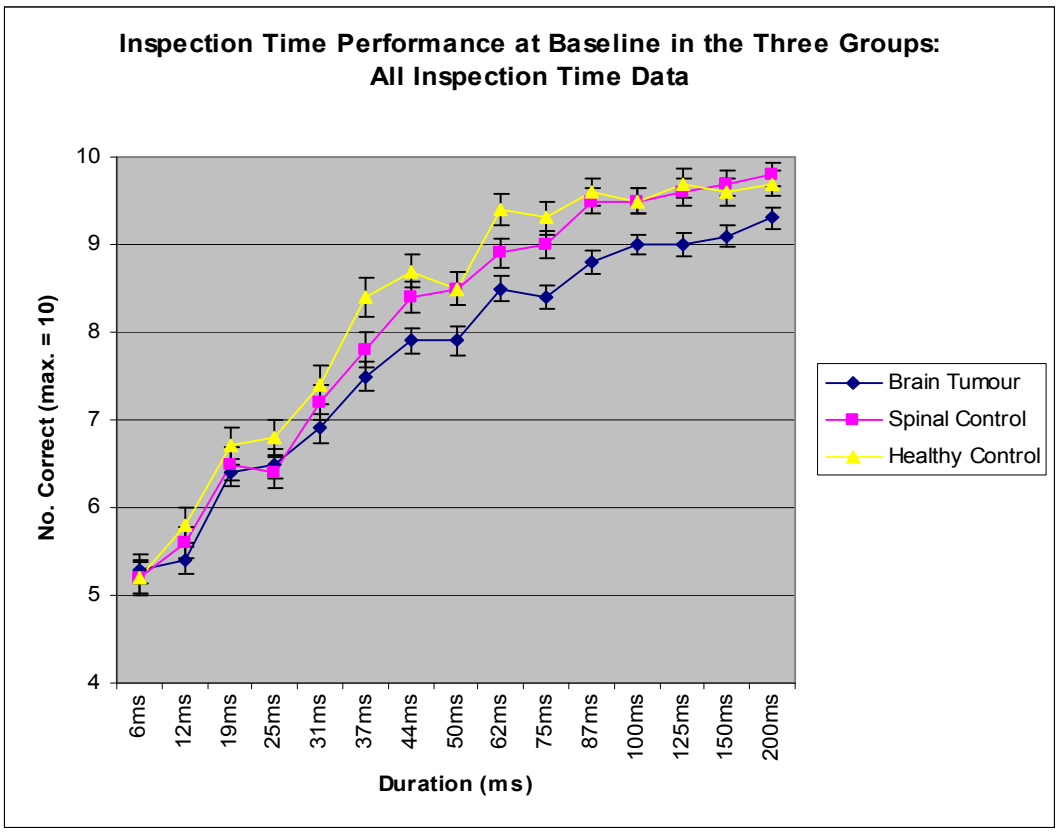


Figure 5.2. Psychometric curves describing inspection time performance (all data) in the brain tumour, spinal surgery and healthy control groups. Points are estimated marginal means adjusted for age and NART score. Error bars: Standard Error Means.

The differences between the brain tumour group and each of the control groups on the inspection time task are demonstrated in Figure 5.2. which shows estimated marginal mean scores, adjusted for age, sex and NART score, out of a possible total of 10 across each duration on the task (i.e. from 6ms – 150ms). The three groups score similarly when the stimuli are shown for very brief time periods. However, at intermediate and the longer durations, the line for the brain tumour patients is shifted to the right, compared with the spinal and healthy control groups. This suggests slower visual information processing in the brain tumour cohort by comparison with both control groups.

5.2.3 Inspection Time Scores: Valid Inspection Time Data

Sixteen brain tumour patients, five spinal surgery controls and two healthy volunteer controls did not perform to criterion on inspection time at baseline and failed to score at least 17/20 on the two longest stimulus durations. These scores were recorded as 'invalid' as the participant might not have fully understood the task instructions (as described in Chapter 2.2). Therefore, the above described-analyses were repeated using only data from those participants with 'valid' inspection time scores.

When the invalid scores were not included in the analysis, the mean baseline inspection time scores were 120.1 (SD 13.7) for the brain tumour group, 123.3 (SD 11.0) for the spinal control group and 127.2 (SD 10.3) for the healthy control group.

The covariates age, $F(1,244) = 73.362$, $p < 0.001$, partial $\eta^2 = 0.23$; and NART score, $F(1,244) = 7.018$, $p = 0.009$, partial $\eta^2 = 0.03$, had a significant main effect in the model that included only valid inspection time scores. Older age and poorer NART scores were significantly associated with lower inspection time scores. The effect of sex was not significant in the model, $F(1,244) = 3.141$, $p = 0.078$, partial $\eta^2 = 0.013$. Participant group had a significant main effect on valid total inspection time scores in the model that included the effect of the covariates age and NART score, $F(2,244) = 6.400$, $p = 0.002$, partial $\eta^2 = 0.05$.

Pairwise comparisons showed a trend towards a difference between the brain tumour and spinal surgery control groups; however, this did not reach the conventional level of statistical significance, ($p = 0.066$). Estimated marginal means, adjusted for age and NART score showed that the brain tumour group scored lower overall than the spinal group on the inspection time task at baseline (120.5 (SE 1.1) vs. 123.4 (SE 1.2), respectively). The difference between the brain tumour group and healthy control group was significant ($p < 0.001$). Estimated marginal means, adjusted for age and NART score, suggest that the brain tumour patients had significantly poorer inspection time scores than the healthy controls (120.5 (SE 1.1) vs. 126.7 (SE 1.3), respectively). The difference between the healthy control group and spinal control group did not reach significance ($p = 0.076$).

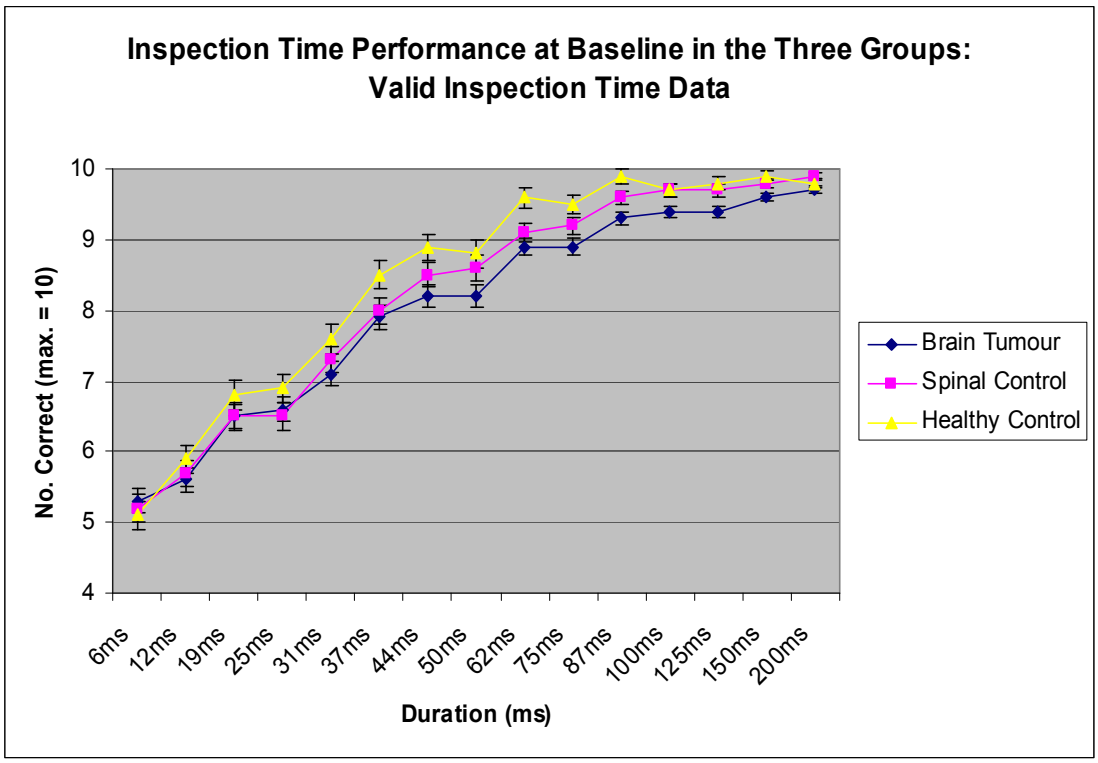


Figure 5.3. Psychometric curves describing inspection time performance (valid data only) in the brain tumour, spinal surgery and healthy control groups. Points are estimated marginal means adjusted for age and NART score. Error bars: Standard Error Means.

Figure 5.3 shows estimated marginal mean scores, adjusted for age and NART score, out of a possible total of 10 for valid inspection time scores, at each of the 15 durations on the inspection time task. Again, the three groups score similarly (i.e. respond at chance) when the stimuli are shown for very brief time periods and also at longer time periods, where the three groups respond almost perfectly. However, at intermediate durations, the line for the brain tumour patients is shifted to the right, when compared with the healthy control group and to a lesser (non-significant) degree when compared with the spinal surgery control group. This again suggests slowed visual information processing in the brain tumour cohort, particularly when compared with the healthy volunteer control group.

5.2.4 Rey Auditory Verbal Learning Test

Eighty-seven brain tumour patients completed the Rey Auditory Verbal Learning Test (RAVLT) at baseline. Seventy-four spinal surgery patients and 80 members of the healthy control group also completed the test.

The mean total number of words recalled across all 8 trials of the RAVLT was 61.7 (SD 17.5) for the brain tumour group; 71.0 words (SD 15.7) for the spinal surgery group and the healthy control group had a mean score of 82.0 words (SD 13.8) words. The number of words correctly recalled was the dependent variable in the general linear model.

There was a significant main effect of the covariates age, $F(1,233) = 84.008$, $p < 0.001$, partial $\eta^2 = 0.27$; and NART score, $F(1,233) = 41.310$, $p < 0.001$, partial $\eta^2 = 0.15$, in the model. Older participants and participants with poorer NART scores were significantly more likely to recall fewer words on the test. Sex also had a significant main effect in the model, $F(1,233) = 6.905$, $p = 0.009$, partial $\eta^2 = 0.03$. Female participants were more likely to have higher scores than male participants. Participant group had a significant main effect in the model that included the effects of age and NART score, $F(2,233) = 18.924$, $p < 0.001$, partial $\eta^2 = 0.14$.

Pairwise comparisons revealed a significant difference between the brain tumour and spinal surgery groups ($p < 0.001$), and a significant difference between the brain tumour and healthy control groups ($p < 0.001$). The estimated marginal mean scores on this test, adjusted for age and NART score, were 64.3 (SE 1.4), 72.2 (SE 1.5) and 77.8 (SE 1.6) for the brain tumour, spinal and healthy control groups, respectively. Therefore, the brain tumour group had significantly lower scores on this memory test than both of the control groups. Additionally, there was a significant difference between the healthy control group and spinal surgery group ($p = 0.014$), showing that the healthy control group significantly outperformed the spinal control group on the RAVLT.

5.2.5 Trail Making Test Part B

The Trail Making Test Part B was completed by 96 members of the brain tumour group, 82 members of the spinal control group and 80 members of the healthy control group.

The mean time taken for the brain tumour group to complete this test was 94.9 (SD 31.7) seconds. Mean scores for the spinal surgery controls and healthy controls were 81.2 (SD 24.1) and 69.0 (SD 20.9) seconds, respectively.

There was a significant main effect of age in the general linear model, $F(1,250) = 56.88$, $p < 0.001$, partial $\eta^2 = 0.19$. The covariate NART score also had a significant main effect in the model, $F(1,250) = 47.72$, $p < 0.001$, partial $\eta^2 = 0.16$. Younger participants and those with lower (better) NART scores were significantly more likely to complete the test in a shorter time. Sex had no significant main effect in the model, $F(1,250) = 0.256$, $p = 0.613$, partial $\eta^2 = 0.001$. There was a significant main effect of group on the test scores in the model that included the effects of age and NART score, $F(2,250) = 10.70$, $p < 0.001$, partial $\eta^2 = 0.08$.

Pairwise comparisons revealed a significant difference between the brain tumour and spinal surgery groups ($p < 0.001$). The estimated marginal mean scores, adjusted for age and NART score, show that the brain tumour group were significantly slower to complete the test than the spinal control group (91.3 secs (SE 2.4) and 79.0 secs (SE 2.6), respectively). The estimated marginal mean score for the healthy control group was 75.2 secs (SE 2.8). The healthy control group were significantly faster than the brain tumour group ($p < 0.001$). The difference between the spinal and healthy control groups was not significant ($p = 0.331$).

5.2.6 Verbal Fluency

Ninety-one participants in the brain tumour group completed the verbal fluency test. Seventy-eight members of the spinal control group and 80 healthy controls also completed the task.

The mean number of words produced on this test was 28.6 (SD 12.0) for the brain tumour group. The spinal control patients produced a mean of 35.0 words (SD 13.1) and healthy controls produced a mean of 45.0 words (SD 12.7) on the task.

There was no significant main effect of age in the model, $F(1,241) = 2.470$, $p = 0.117$, partial $\eta^2 = 0.010$. The covariate NART score did have a significant main effect in the model, $F(1,241) = 70.56$, $p < 0.001$, partial $\eta^2 = 0.23$. Sex also had a significant main effect, $F(1,241) = 5.94$, $p = 0.016$, partial $\eta^2 = 0.02$. Female participants and participants with lower (better) NART scores were significantly more likely the have better scores on the verbal fluency test. There was a significant main effect of group on verbal fluency scores in the model that included the effects of the covariates, $F(2,241) = 13.77$, $p < 0.001$, partial $\eta^2 = 0.10$.

Pairwise comparisons using LSD tests revealed a significant difference between the brain tumour group and the spinal surgery control group ($p = 0.002$). The estimated marginal mean score, adjusted for age and NART score, for the brain tumour group was 31.0 (SE 1.2) on this test, and for the spinal control group was 36.4 (SE 1.3). Thus, the brain tumour group produced significantly fewer words than their spinal surgery counterparts. LSD tests also revealed a significant difference between the brain tumour and healthy control groups ($p < 0.001$). The estimated marginal mean score for the healthy control group was 40.8 (SE 1.4) showing that the brain tumour group also had significantly lower scores than the healthy control group. The difference between the spinal and healthy control groups was significant ($p = 0.023$). The healthy control group performed better than the spinal surgery control group on the verbal fluency task at baseline.

5.2.7 Digit Symbol Coding

One hundred and eleven brain tumour patients, 83 spinal controls and 79 healthy controls completed the digit symbol-coding task at baseline.

Mean scores on this test were 56.4 (SD 22.2), 67.4 (SD 16.9) and 76.2 (SD 15.7) for the brain tumour, spinal control and healthy control groups, respectively.

There was a significant main effect of the covariates age, $F(1,265) = 113.59$, $p < 0.001$, partial $\eta^2 = 0.30$; and NART score, $F(1,265) = 57.72$, $p < 0.001$, partial $\eta^2 = 0.18$, in the model. Older participants and participants who performed less well on the NART were significantly more likely to have lower digit symbol-coding scores. There was no significant main effect of sex on digit symbol coding scores, $F(1,265) = 2.929$, $p = 0.088$, partial $\eta^2 = 0.011$. In the model that included the effects of the covariates NART and age, there was a significant main effect of group, $F(2,265) = 13.24$, $p < 0.001$, partial $\eta^2 = 0.09$.

The estimated marginal mean scores, adjusted for age and NART score, were 59.3 (SE 1.5) for the brain tumour group, 68.0 (SE 1.7) for the spinal control group and 70.9 (SE 1.9) for the healthy control group. Pairwise comparisons using LSD tests revealed that the brain tumour group performed significantly less well than the spinal surgery group ($p < 0.001$). The brain tumour group also performed significant less well than the healthy control group ($p < 0.001$). The difference between the spinal and healthy control groups was not significant ($p = 0.266$).

5.2.8 Letter-Number Sequencing

The Letter-Number Sequencing test was administered to 89 brain tumour patients, 78 spinal control patients and 79 healthy control participants.

The brain tumour group scored a mean of 9.4 (SD 3.4) on the test, the spinal control group scored a mean of 11.2 (SD 2.7) and the healthy controls had a mean score of 11.2 (SD 2.6).

There was a significant main effect of age on task performance, $F(1,238) = 69.05$, $p < 0.001$, partial $\eta^2 = 0.23$. The covariate NART score also had a significant main effect in the model, $F(1,238) = 59.34$, $p < 0.001$, partial $\eta^2 = 0.20$. Older participants and those participants with higher (poorer) NART scores were significantly more likely to have lower scores on this test. Sex had no significant main effect in the model, $F(1,238) = 0.379$, $p = 0.539$, partial $\eta^2 = 0.002$. Participant group had a

significant main effect in the model that included the effects of the covariates age and NART score, $F(2, 238) = 9.67$, $p < 0.001$, partial $\eta^2 = 0.08$.

The estimated marginal mean scores, adjusted for age and NART score, were 9.9 (SE 0.3) for the brain tumour group, 11.5 (SE 0.3) for the spinal surgery control group and 10.3 (SE 0.3) for the healthy control group. Pairwise comparisons revealed that the brain tumour group performed significantly less well than the spinal surgery patients ($p < 0.001$). However, there was no significant difference between the brain tumour patients and the healthy control group ($p = 0.335$). Although the estimated marginal mean score for the healthy control group was higher than that of the brain tumour group, the difference was not statistically significant. The performance of the spinal and healthy control groups were also significant different ($p = 0.004$). The spinal surgery control group performed significantly better than the healthy control group on this occasion.

5.3 Baseline Analyses: Functional Measures

5.3.1 Overview of Analysis Procedure

The Williams Delayed Recall Test (WDRT), Nine Hole Peg Test (NPHT) and Timed Ten Metre Walk (TTMW) subtests of the Edinburgh Functional Impairment Tests (EFITs) were analysed in the same way as the cognitive test scores described in section 5.2.1, using general linear modelling (analysis of covariance). Age and NART were covariates and group (brain tumour, spinal or healthy) and sex were fixed factors in each model.

5.3.2 Williams Delayed Recall Test (EFIT)

One hundred and fifteen brain tumour patients completed this EFIT subtest. Eighty-three spinal surgery patients and 80 healthy control participants also completed the WDRT.

The brain tumour group had a mean score of 8.2 (SD 5.1) on the WDRT. The mean score for the spinal control group was 5.3 (SD 3.8) and was 4.5 (SD 3.6) for the healthy control group.

There was a significant main effect of both of the covariates age, $F(1,270) = 38.24$, $p < 0.001$, partial $\eta^2 = 0.12$; and NART score, $F(1,270) = 11.01$, $p = 0.001$, partial $\eta^2 = 0.04$ in the model. Younger participants and participants with lower (better) NART scores were significantly more likely to have lower (better) WDRT scores. Sex had no significant main effect in the model, $F(1,270) = 2.527$, $p = 0.113$, partial $\eta^2 = 0.009$. The effect of group was significant in the model that included the effects of age and NART score, $F(2,270) = 12.34$, $p < 0.001$, partial $\eta^2 = 0.08$.

Estimated marginal mean scores and pairwise comparisons using LSD tests indicated that the brain tumour group had significantly higher scores than the spinal control group (7.8 (SE 0.4) vs. 5.3 (SE 0.4), with higher scores on this test indicating poorer performance, $p < 0.001$). The difference between the brain tumour and the healthy control group was also significantly different (7.8 (SE 0.4) vs. 5.2 (SE 0.5), $p < 0.001$). This suggests that the brain tumour patients had significantly higher scores (i.e. poorer performance) on this test at baseline than either of the control groups. The difference between the healthy control group and the spinal control group was not significant ($p = 0.855$).

5.3.3 Nine Hole Peg Test (Right Hand, EFIT)

In the brain tumour group, 115 patients completed this test using their right hand. Eighty-three spinal controls and 80 healthy controls also completed the test.

For this test, the dependent variable was the time taken to complete the test using the right hand, in seconds. The mean time taken to complete the test was 14.7 seconds (SD 4.5) for the brain tumour group, 12.9 seconds (SD 2.7) for the spinal control group and 12.3 seconds (SD 1.5) for the healthy control group.

There was a significant main effect of the covariates age, $F(1,270) = 33.44$, $p < 0.001$, partial $\eta^2 = 0.11$; and NART score, $F(1,270) = 19.47$, $p < 0.001$, partial $\eta^2 = 0.06$ in the model. This suggests that older age and higher (poorer) NART scores were significantly associated with taking longer to complete the test. There was no significant main effect of sex in the model, $F(1,270) = 1.829$, $p = 0.117$, partial $\eta^2 = 0.007$. Group had a significant main effect in the model that included the effects of age and premorbid ability (NART score), $F(2,270) = 6.55$, $p = 0.002$, partial $\eta^2 = 0.05$.

The estimated marginal mean scores, adjusted for age and NART score, were 14.4 (SE 0.3) for the brain tumour group, 12.9 (SE 0.4) for the spinal group and 13.0 (SE 0.4) for the healthy control group. Pairwise comparisons using LSD tests revealed a significant difference between the performance of the brain tumour and spinal surgery groups ($p = 0.001$) and between the brain tumour and healthy control groups ($p = 0.008$). Therefore, the brain tumour group were significantly slower than each of the control groups on the NHPT with the right hand. The difference between the two control groups (spinal surgery and healthy volunteer) was not significant ($p = 0.789$).

5.3.4 Nine Hole Peg Test (Left Hand, EFIT)

This test was completed using the left hand by 113 brain tumour patients, 83 spinal surgery controls and 80 healthy controls.

Mean scores (i.e. time taken to complete the test using the left hand) for this test were 16.6 seconds (SD 13.5) for the brain tumour group, 14.3 seconds (SD 1.6) for the spinal control group and 12.8 (SD 1.9) for the healthy control group.

In contrast with the model for the right-hand NHPT, there was only a significant main effect of the covariate age on the time taken to complete the test, $F(1,268) = 12.00$, $p = 0.001$, partial $\eta^2 = 0.04$. Older participants took significantly longer to complete the test. NART score had no significant main effect in the model, $F(1,268) = 0.016$, $p = 0.898$, partial $\eta^2 = 0.000$; nor did sex, $F(1,268) = 0.380$, $p = 0.538$, partial $\eta^2 = 0.001$. Participant group had a significant main effect in the model that

included the effects of age and NART score, $F(2,268) = 3.74$, $p = 0.025$, partial $\eta^2 = 0.03$.

The estimated marginal mean scores on this test, adjusted for age and NART score, were 16.5 (SE 0.9) for the brain tumour group, 14.5 (SE1.0) for the spinal control group, and 12.7 (SE 1.1) for the healthy control group. The difference between the brain tumour and healthy control groups was significant ($p = 0.008$), showing that the brain tumour patients too significantly longer to complete the task than the healthy control group. There were no significant differences between the brain tumour and spinal surgery groups ($p = 0.111$); or between the healthy and spinal surgery control groups ($p = 0.237$).

5.3.5 Timed Ten Metre Walk (EFIT)

One hundred and two brain tumour patients, 77 spinal controls and 77 healthy controls completed this test.

The dependent variable for this measure was the time take to complete the test, in seconds. The brain tumour group took a mean of 7.0 seconds (SD 1.8) to complete this task. Mean scores for the spinal and healthy control groups were 8.0 seconds (SD 2.3) and 6.1 seconds (SD 1.2), respectively.

There was a significant main effect of the covariate age in this model, $F(1,248) = 33.85$, $p < 0.001$, partial $\eta^2 = 0.12$. Older participants were slower to complete the TMW. Neither NART score nor sex had a significant main effect in the model, $F(1,248) = 3.647$, $p = 0.057$, partial $\eta^2 = 0.014$; and $F(1,248) = 3.113$, $p = 0.079$, partial $\eta^2 = 0.012$, respectively. Participant group had a significant main effect in the model that included the effects of the covariates, $F(2,248) = 19.38$, $p < 0.001$, partial $\eta^2 = 0.14$.

The estimated marginal mean scores, adjusted for age and NART score, were 6.9 (SE 0.2) for the brain tumour group, 8.0 (SE 0.2) for the spinal group and 6.2 (SE 0.2) for the healthy control group. Pairwise comparisons using LSD tests revealed that the

brain tumour group were significantly faster than the spinal surgery group ($p < 0.001$). The brain tumour group were, however, significantly slower than the healthy controls ($p = 0.012$). The spinal control group were also significantly slower than the healthy control group ($p < 0.001$).

5.4 Baseline Analyses: Mood

The Hospital Anxiety and Depression Scale (HADS) was administered as a measure of mood and was completed by 115 brain tumour patients, 82 spinal controls and 80 healthy control participants at baseline. In the total sample of the three participant groups combined, NART score was significantly correlated with total HADS score ($r(n=277) = 0.226$; $p < 0.001$). Therefore, NART score was also included as a covariate in the following analyses.

5.4.1 Anxiety Scores

The mean anxiety scores for each group on this questionnaire were 8.0 (SD 4.4) for the brain tumour group, 7.9 (SD 3.8) for the spinal control group and 6.0 (SD 3.6) for the healthy control group.

General linear modelling revealed no significant main effect of the covariate age, $F(1,269) = 0.011$, $p = 0.918$, partial $\eta^2 = 0.000$; the covariate NART score, $F(1,269) = 1.164$, $p=0.282$, partial $\eta^2 = 0.004$; or of sex, $F(1,269) = 3.175$, $p = 0.076$, partial $\eta^2 = 0.012$, in the model. There was however a significant main effect of participant group in the model that included the effects of the covariates, $F(2,269) = 4.76$, $p = 0.009$, partial $\eta^2 = 0.03$.

Pairwise comparisons showed that the brain tumour group and healthy control group scores were significantly different ($p = 0.004$). The estimated marginal mean score for the brain tumour group was 7.9 (SE 0.4), compared with 6.0 (SE 0.5) for the healthy control group, showing that the brain tumour group had significantly higher anxiety scores than the healthy control group. There was no significant difference between the brain tumour and spinal control group ($p = 0.835$). The estimated

marginal mean score for the spinal cohort was 7.8 (SE 0.4). The spinal group had significantly higher anxiety scores than the healthy controls ($p = 0.010$).

5.4.2 Depression Scores

The mean scores on the depression scale of the HADS were 4.7 (SD 4.1) for the brain tumour group; 5.3 (SD 3.6) for the spinal control group and 2.5 (SD 2.1) for the healthy control group.

The covariate age had no significant main effect in the model, $F(1,269) = 2.934$, $p = 0.088$, partial $\eta^2 = 0.011$; nor did sex, $F(1,269) = 1.962$, $p = 0.162$, partial $\eta^2 = 0.007$. The covariate NART score did have a significant main effect in the model, $F(1,269) = 6.69$, $p = 0.010$, partial $\eta^2 = 0.02$. Participant group had a significant main effect in the model that included the effects of the covariates, $F(2,269) = 8.03$, $p < 0.001$, partial $\eta^2 = 0.06$.

Estimated marginal mean depression scores, adjusted for age and NART score, were 4.5 (SE 0.3) for the brain tumour group, 5.2 (SE 0.4) for the spinal control group and 2.9 (SE 0.4) for the healthy control groups. Pairwise comparisons showed that the brain tumour group had significantly higher depression scores than the healthy control group ($p = 0.004$). However, there was no significant difference between the depression scores in the brain tumour and spinal surgery groups ($p = 0.174$). The spinal control group had significantly higher levels of depression as measured by the HADS than the healthy controls ($p < 0.001$).

5.4.3 Total Hospital Anxiety and Depression Scale Scores

Mean total score (i.e. anxiety and depression scores added together) on the Hospital Anxiety and Depression Scale for the brain tumour group was 12.7 (SD 7.5). The spinal group had a total score mean of 13.2 (SD 6.5) and the healthy control group had a mean of 8.5 (SD 5.0).

The covariate NART score had a significant main effect in the model, $F(1,269) = 4.05$, $p = 0.045$, partial $\eta^2 = 0.02$. There was no significant main effect of the

covariate age, F (1,269) = 0.919, p = 0.339, partial η^2 = 0.003; or sex, F(1,269) = 3.316, p = 0.070, partial η^2 = 0.012, in the model. There was a significant main effect of group on total HADS scores in the model that included the effects of the covariates, F(2,269) = 7.59, p = 0.001, partial η^2 = 0.05.

The estimated marginal mean total scores on this questionnaire, adjusted for age and NART score, were 12.4 (SE 0.6) for the brain tumour group, 13.0 (SE 0.7) for the spinal control group and 8.9 (SE 0.8) for the healthy control group. Pairwise comparisons revealed a significant difference between the brain tumour and healthy control groups (p = 0.001), suggesting that the brain tumour group had higher scores on the HADS than the healthy control participants. There was no significant difference between the brain tumour and spinal surgery group scores (p 0.559). The spinal surgery control group had significantly higher scores than the healthy control group (p < 0.001).

5.5 Baseline Analyses: Functional Measures

5.5.1 Overview of Analysis Procedure

Since the Barthel Disability Index, Karnofsky Performance Scale and Boston Aphasia Severity Rating Scale were only assessed for the brain tumour and spinal groups, the same analyses as detailed above were carried out but with brain tumour and spinal surgery patients comprising the two groups in the model. However, since sex had no significant effect in any of the models for these functional scales it was therefore excluded as a fixed factor in the analyses reported below.

5.5.2 Barthel Disability Index

The brain tumour patients had a mean score of 19.6 (SD 1.3) out of a possible score of 20, on the Barthel Disability Index, suggesting that the majority of patients had

little or no disability. The spinal surgery patients scored a mean of 19.4 (SD 1.2) on the measure.

There was a significant main effect of the covariate age in the model, $F(1,195) = 10.514$, $p = 0.001$, partial $\eta^2 = 0.051$. There was also a significant main effect of NART score on Barthel Disability Index scores, $F(1,195) = 3.975$, $p = 0.048$, partial $\eta^2 = 0.020$. However, there was no significant main effect of participant group in the model that included the effects of the covariates, $F(1,195) = 1.289$, $p = 0.258$, partial $\eta^2 = 0.007$. This suggests that the brain tumour and spinal surgery groups did not differ significantly in terms of disability as measured by the Barthel Disability Index.

5.5.3 Karnofsky Performance Scale

The brain tumour group had a mean score of 87.9 out of a possible 100 (SD 10.2) on this performance scale, compared with a mean score of 80.4 (SD 8.2) for the spinal surgery group.

There was a significant main effect of the covariates age, $F(1,195) = 23.775$, $p < 0.001$, partial $\eta^2 = 0.109$; and NART score, $F(1,195) = 18.246$, $p < 0.001$, partial $\eta^2 = 0.086$ in the model. Participant group had a significant main effect in the model that included the effects of age and NART score, $F(1,195) = 43.972$, $p < 0.001$, partial $\eta^2 = 0.184$. Estimated marginal mean scores, adjusted for age and NART score, were 88.2 (SE 0.8) for the brain tumour group and 80.0 (SE 0.1) for the spinal control group, suggesting spinal patients had lower scores (indicating a poorer performance status) on the Karnofsky Performance Scale.

5.5.4 Boston Aphasia Severity Rating Scale (EFIT)

A mean rating of 5.9 (SD 0.4) of a possible 6 was given to the brain tumour patients. The spinal patients all had ‘normal’ speech ratings, with a mean score of 6 (SD 0.0).

There was no significant main effect of either of the covariates age, $F(1,194) = 3.458$, $p = 0.064$, partial $\eta^2 = 0.018$; or NART score, $F(1,194) = 0.158$, $p = 0.692$, partial $\eta^2 = 0.001$, in the model. There was, however, a significant main effect of participant group in the model, $F(1,194) = 4.100$, $p = 0.044$, partial $\eta^2 = 0.021$. Estimated marginal means, adjusted for age and sex, suggest that the brain tumour patients had significantly lower ratings on the scale than the spinal surgery patients (5.9 (SE 0.0) vs. 6.0 (SE 0.0)).

Table 5.1. Demographic characteristics of each study group: brain tumour, spinal surgery and healthy control.

	Brain Tumour Group (n = 118)	**Spinal Control Group (n = 85)**	**Healthy Control Group (n = 80)**	**p value for difference***
Sex	60 male, 58 female	39 male, 46 female	31 male, 49 female	0.245
Mean Age (SD) (Range)	50.02 (12.98) (17-80)	47.76 (11.23) (23 – 80)	49.05 (16.38) (22 – 79)	0.507
Educational Background				
- No formal education	17	9	0	< 0.001
- O'Grades (or equivalent)	47	19	5	
- Highers (or equivalent)	18	12	7	
- Teaching Qualification/HND/HNC	14	24	6	
- Degree	20	19	59	
- Other	1	2	0	
- Unknown	1	0	0	
Handedness	103 right, 15 left	78 right, 6 left, 1 ambidextrous	74 right, 6 left	0.318

* Based on Pearson Chi-square tests, except for age (One-way ANOVA)

Table 5.2. Crosstabulation showing actual vs. expected counts for qualifications held in each of the three groups (brain tumour, spinal surgery, healthy)

			Highest Qualification Held						
			No formal qualifications	O' Grades or equivalent	CSYS/ Highers or equivalent	Teaching qualification / HND/ HNC	Degree or equivalent	Other	Unknown
Group	Brain tumour	Actual Count	17	47	18	14	20	1	1
		Expected Count	10.8	29.6	15.4	18.8	41.7	1.3	0.4
	Spinal	Actual Count	9	19	12	24	19	2	0
		Expected Count	7.8	21.3	11.1	13.5	30	0.9	0.3
	Healthy	Actual Count	0	5	7	7	61	0	0
		Expected Count	7.3	20.1	10.5	12.7	28.3	0.8	0.3

Table 5.3. Histological and location characteristics of the brain tumour group.

		n	Percentage
Histological subtype	**High – Grade Glioma (WHO III – IV)**	51	43.2
	Low – Grade Glioma (WHO I – II)	23	19.5
	Metastasis	14	11.9
	Meningioma	17	14.4
	Other	11	9.3
	Unknown (no surgery)	2	1.7
Hemispheric Location	**Left**	56	47.5
	Right	56	47.5
	Bi-hemispheric	2	1.7
	Not applicable	4	3.4
Lobe	**Frontal**	34	28.8
	Temporal	16	13.6
	Limbic	8	6.8
	Parietal	19	16.1
	Occipital	5	4.2
	> 1 Lobe	23	19.5
	Other	13	11.0

Table 5.4. Overview of results of general linear modelling analyses comparing the brain tumour, spinal surgery and healthy control groups on cognitive and mood scales.

	Group	Sex	N	Mean[3]	Standard Deviation	Effect of Group			Effect of Age		
						F	Sig. (p)	n^{2}*	F	Sig. (p)	n^{2}*
Inspection Time (Total, all data) [1]	Brain Tumour	Male	56	118.7	19.2	7.994	< 0.001	0.057	76.946	<0.001	0.224
		Female	56	109.9	19.7						
		Total	112	114.3	19.9						
	Spinal	Male	38	121.6	13.5						
		Female	44	121.2	14.0						
		Total	82	121.4	13.7						
	Healthy	Male	31	126.1	13.1						
		Female	49	126.5	10.6						
		Total	80	126.4	11.5						
Inspection Time (Total, valid data)[1]	Brain Tumour	Male	50	123.4	11.9	6.400	0.002	0.050	73.362	<0.001	0.231
		Female	46	116.5	14.6						
		Total	96	120.1	13.7						
	Spinal	Male	36	123.5	11.0						
		Female	42	123.1	11.1						
		Total	78	123.3	11.0						
	Healthy	Male	30	127.3	11.5						
		Female	48	125.2	9.7						
		Total	78	127.2	10.3						

Rey Auditory Verbal Learning Test (Total) [1]	Brain Tumour	Male	46	59.5	17.1	18.924	<0.001	0.140	84.008	<0.001	0.265
		Female	41	64.1	17.2						
		Total	87	61.7	17.5						
	Spinal	Male	34	71.0	14.5						
		Female	40	71.1	16.9						
		Total	74	71.0	15.7						
	Healthy	Male	31	77.6	14.6						
		Female	49	84.8	12.5						
		Total	80	82.0	13.8						
Trail Making Test Part B (secs) [2]	Brain Tumour	Male	51	93.2	29.3	10.703	<0.001	0.079	56.881	<0.001	0.185
		Female	45	96.7	34.5						
		Total	96	94.9	31.7						
	Spinal	Male	37	81.0	27.8						
		Female	45	81.4	20.8						
		Total	82	81.2	24.1						
	Healthy	Male	31	68.3	21.0						
		Female	49	69.5	20.9						
		Total	80	69.0	20.8						

Verbal Fluency (Total)[1]	Brain Tumour	Male	45	26.7	9.4	13.772	<0.001	0.103	2.470	0.117	0.010
		Female	46	30.6	14.0						
		Total	91	28.7	12.0						
	Spinal	Male	38	32.2	13.5						
		Female	40	37.6	12.3						
		Total	78	35.0	13.1						
	Healthy	Male	31	44.7	13.8						
		Female	49	45.2	12.0						
		Total	80	45.0	12.7						
Digit Symbol Coding (Total)[1]	Brain Tumour	Male	55	57.9	19.7	13.236	<0.001	0.091	113.58	<0.001	0.300
		Female	56	54.9	24.5						
		Total	111	56.4	22.2						
	Spinal	Male	38	65.1	16.8						
		Female	45	69.3	17.0						
		Total	83	67.4	16.9						
	Healthy	Male	31	72.7	14.6						
		Female	48	78.5	16.1						
		Total	79	76.2	15.7						

Letter-Number Sequencing (Total)[1]	Brain Tumour	Male	45	9.4	3.1	9.674	<0.001	0.075	69.054	<0.001	0.225
		Female	44	9.4	3.6						
		Total	89	9.4	3.4						
	Spinal	Male	36	11.5	2.5						
		Female	42	10.9	2.8						
		Total	78	11.2	2.7						
	Healthy	Male	30	10.5	2.4						
		Female	49	11.6	2.7						
		Total	79	11.2	2.6						
EFIT Williams Delayed Recall Test (Total)[2]	Brain Tumour	Male	58	8.4	5.2	12.342	<0.001	0.084	38.241	<0.001	0.124
		Female	57	8.0	5.0						
		Total	115	8.2	5.1						
	Spinal	Male	38	5.7	4.1						
		Female	45	5.0	3.5						
		Total	83	5.3	3.8						
	Healthy	Male	31	5.2	4.1						
		Female	49	4.1	3.2						
		Total	80	4.5	3.6						

EFIT Nine Hole Peg Test (Right Hand, secs)²	Brain Tumour	Male	58	14.8	4.1	6.547	0.002	0.046	33.441	<0.001	0.110
		Female	57	14.7	5.0						
		Total	115	14.7	4.5						
	Spinal	Male	38	13.2	2.0						
		Female	45	12.7	3.1						
		Total	83	12.9	2.7						
	Healthy	Male	31	12.8	1.5						
		Female	49	12.0	1.5						
		Total	80	12.3	1.5						
EFIT Nine Hole Peg Test (Left Hand, secs)²	Brain Tumour	Male	56	15.6	3.5	3.742	0.025	0.027	11.997	0.001	0.043
		Female	57	17.6	18.8						
		Total	113	16.6	13.5						
	Spinal	Male	38	14.1	2.4						
		Female	45	14.5	4.4						
		Total	83	14.3	3.6						
	Healthy	Male	31	13.1	1.9						
		Female	49	12.7	1.9						
		Total	80	12.8	1.9						

EFIT Timed Ten Metre Walk (secs)²	Brain Tumour	Male	55	6.7	1.6	19.377	<0.001	0.135	33.845	<0.001	0.120
		Female	47	7.3	1.9						
		Total	102	7.0	1.8						
	Spinal	Male	34	7.6	2.1						
		Female	43	8.3	2.5						
		Total	77	8.0	2.3						
	Healthy	Male	31	6.1	1.2						
		Female	46	6.0	1.3						
		Total	77	6.1	1.2						
Hospital Anxiety and Depression Scale – Anxiety Score²	Brain Tumour	Male	58	7.6	4.3	4.764	0.009	0.034	0.011	0.918	<0.001
		Female	57	8.4	4.4						
		Total	115	8.0	4.4						
	Spinal	Male	38	7.7	4.2						
		Female	44	8.1	3.5						
		Total	82	7.9	3.8						
	Healthy	Male	31	5.1	3.2						
		Female	49	6.6	3.7						
		Total	80	6.0	3.5						

Hospital Anxiety and Depression Scale – Depression Score²	Brain Tumour	Male	58	4.1	3.6	8.028	<0.001	0.056	2.934	0.088	0.011
		Female	57	5.4	4.4						
		Total	115	4.7	4.1						
	Spinal	Male	38	4.8	3.3						
		Female	44	5.8	3.8						
		Total	82	5.3	3.6						
	Healthy	Male	31	2.7	2.1						
		Female	49	2.3	2.2						
		Total	80	2.5	2.1						
Hospital Anxiety and Depression Scale –Total Score²	Brain Tumour	Male	58	11.7	7.1	7.589	0.001	0.053	0.919	0.339	0.003
		Female	57	13.8	7.9						
		Total	115	12.7	7.5						
	Spinal	Male	38	12.4	6.6						
		Female	44	13.8	6.3						
		Total	82	13.2	6.5						
	Healthy	Male	31	7.8	4.6						
		Female	49	8.9	5.2						
		Total	80	8.5	5.0						

[1] Higher scores represent better function
[2] Higher scores represent worse function
[3] Raw mean score
* n^2 = the proportion of variance accounted for by the covariate (group or age)

Table 5.5. Estimated marginal mean scores, adjusted for age and NART score, for each participant group on each baseline test and post-hoc pairwise comparisons.

	Estimated Marginal Mean (Standard Error)			Pairwise Comparisons – p-value for significance	
	Brain Tumour	**Spinal Surgery**	**Healthy Control**	**Brain Tumour vs. Spinal Surgery**	**Brain Tumour vs. Healthy Control**
Inspection Time (Total, all data) [1]	115.9 (1.3)	121.6 (1.6)	124.2 (1.7)	0.005	< 0.001
Inspection Time (Total, valid data)[1]	120.5 (1.1)	123.4 (1.2)	126.7 (1.3)	0.066	< 0.001
Rey Auditory Verbal Learning Test (Total) [1]	64.3 (1.4)	72.2 (1.5)	77.8 (1.6)	< 0.001	< 0.001
Trail Making Test Part B (secs) [2]	91.3 (2.4)	79.0 (2.6)	75.2 (2.8)	< 0.001	< 0.001
Verbal Fluency (Total)[1]	31.0 (1.2)	36.4 (1.3)	40.8 (1.4)	0.002	< 0.001
Digit Symbol Coding (Total) [1]	59.3 (1.5)	68.0 (1.7)	70.9 (1.9)	< 0.001	< 0.001

Letter-Number Sequencing (Total)[1]	9.9 (0.3)	11.5 (0.3)	10.3 (0.3)	< 0.001	0.335
EFIT Williams Delayed Recall Test (total)[2]	7.8 (0.4)	5.3 (0.4)	5.2 (0.5)	< 0.001	< 0.001
EFIT Nine Hole Peg Test (Right Hand, secs)[2]	14.4 (0.3)	12.9 (0.4)	13.0 (0.4)	0.001	0.008
EFIT Nine Hole Peg Test (Left Hand, secs)[2]	16.5 (0.9)	14.5 (1.0)	12.7 (1.1)	0.111	0.008
EFIT Timed Ten Metre Walk (secs)[2]	6.9 (0.2)	8.0 (0.2)	6.2 (0.2)	< 0.001	0.012
Hospital Anxiety and Depression Scale – Anxiety Score[2]	7.9 (0.4)	7.8 (0.4)	6.0 (0.5)	0.835	0.004
Hospital Anxiety and Depression Scale – Depression Score[2]	4.5 (0.3)	5.2 (0.4)	2.9 (0.4)	0.174	0.004
Hospital Anxiety and Depression Scale –Total Score[2]	12.4 (0.6)	13.0 (0.7)	8.9 (0.8)	0.559	0.001

[1] Higher scores represent better function

[2] Higher scores represent worse function

5.6 Discussion

There was a significant overall effect of participant group (brain tumour, spinal surgery control or healthy control) on each of the measures administered at baseline. The brain tumour group were found to perform significantly worse than both the spinal surgery and healthy control groups on the majority of tests administered. Visual information processing, as measured by the inspection time task, was significantly disrupted in brain tumour patients by comparison with both control groups and the effect size was moderate. When 'invalid' inspection time scores were excluded a significant impairment, again of moderate effect size, was found in the brain tumour group by comparison with the healthy control group. The difference between the brain tumour group and the spinal surgery group did not reach the conventional level of statistical significance when only valid inspection time data was included, although there was a clear trend towards poorer performance in the brain tumour group. The brain tumour group performed significantly less well than both the spinal surgery and healthy control groups on most of the other cognitive measures at baseline. The effect size of participant group was large in the model for the Rey Auditory Verbal Learning Test and was moderate for the majority of the other cognitive tests. The spinal surgery group performed significantly worse than both the brain tumour group and healthy control group on the Timed Ten Metre Walk and the brain tumour group, in turn, were significantly slower than the healthy control group on this measure. Both the brain tumour and the spinal surgery groups had significantly higher levels of anxiety and depression than the healthy control group, as measured by the HADS. However, HADS scores did not differ significantly between the brain tumour and spinal surgery groups. Disability, as measured by the Barthel Index did not differ between the two surgical groups, although the spinal surgery group had significantly poorer performance status than the brain tumour group, as measured by the Karnofsky Performance Scale. Thus, these findings suggest that patients with primary and secondary brain tumours located throughout the brain have significantly impaired cognition, at the time of presentation, prior to surgery by comparison with matched surgical and healthy volunteer controls.

One of the main aims of this thesis was to evaluate the utility of the inspection time task for use in a neuro-oncological setting. The inspection time task proved to be a feasible and useful measure of visual information processing that was sensitive to impaired function in the brain tumour cohort at the time of presentation to a neurosurgical department, prior to surgery. There was a moderate effect size of group in the inspection time model. However, the effect sizes for the Rey Auditory Verbal Learning Test, Verbal Fluency and Digit Symbol Coding in particular, were large and this suggests that inspection time, although evidently able to detect a difference between the groups, was perhaps not as sensitive a measure as the aforementioned tests. However, inspection time was equally as sensitive as a number of other commonly-used standardised tests, suggesting it may have a potential role for inclusion in cognitive test batteries in this patient group. The results of the present study do however differ from those of a pilot study that tested 23 brain tumour patients and 24 spinal surgery controls in the pre and post-operative period on inspection time and other cognitive measures (Zbinden et al., 2006). The preliminary investigation found that patients in the brain tumour cohort had significantly lower inspection time scores than the spinal surgery control group in the pre-operative period, a finding that is confirmed by the present study. However, in the pilot study, the brain tumour group performed equally as well as the spinal surgery control group on a number of other standardised cognitive tests administered pre-operatively, including the Rey Auditory Verbal Learning Test, Digit Symbol Coding, the Trail Making Test Part B and the Edinburgh Functional Impairment Tests. Conversely, in the present study, the brain tumour group were also significantly impaired on these tests prior to surgery. Since only 12 of 23 brain tumour patients recruited into the pilot study completed these other cognitive measures and it may be that those patients who failed to complete the test battery in its entirety were those patients who were most severely affected by their disease. Performing a large battery of psychometric and cognitive tests can be difficult for brain tumour patients, particularly those who experience neurological symptoms. Indeed, as highlighted previously, a number of patients in the brain tumour cohort in the present study also failed to complete a number of the cognitive tests for a variety of reasons including

poor comprehension of the task, fatigue or motor impairment that prevented task completion (e.g. on the trail making test part B). However, the larger cohort recruited into the present study may have had more power to detect impairment across a number of cognitive domains.

Thus, inspection time has proven to be a useful measure that has many practical advantages as a tool in clinical neuro-oncology. Namely the task does not require intact motor function as patients can voice their responses and conversely, it does not require intact speech if the patient is able to respond independently. The large number of practice trials also give patients who may be mildly confused the opportunity to familiarise themselves with the task demands prior to the test trials. The inspection time task was well tolerated by the majority of patients enrolled into the study, with only 4 of the 118 brain tumour patients (3%) failing to complete the task in its entirety at baseline testing. However, 14% of those brain tumour patients who did complete the task did not achieve 'valid' inspection time scores at baseline (i.e. a score of ≥ 17/20 on the longest two durations). Only 6% of the spinal surgery group and 3% of the healthy control group had 'invalid' scores at baseline. This validity criterion has been applied in previous inspection time studies because the longest two stimulus presentation durations (150ms and 200ms) are easily visible to the majority of participants and therefore, a near perfect score should be obtained on these longest trials. Thus, any participant who scores less than 17/20 at these longer durations may not have fully understood the task. However, the question of whether such validity criterion can be applied to a cohort of patients with brain tumours has been raised following the present study. The majority of patients who did not meet validity criterion were noted to have appeared to understand the task during the practice trials. 'Valid' scores may have been obtained by some of these patients had the stimuli been shown for longer durations. It is possible that, in some brain tumour patients, visual information processing is slowed to such an extent that invalid inspection time scores reflect severe impairment, as opposed to lack of comprehension of the task. If this is the case, a revised version of the inspection time task that involves presenting the stimulus for durations longer than 200ms would provide further insight into the extent of visual information processing impairment in

neuro-oncological patients. Moreover, given that the inspection time task can take up to 30 minutes to complete, it therefore has relatively high attentional demands, particularly by comparison with other, shorter cognitive tests. Thus, impaired concentration could also explain, in part, the high proportion of 'invalid' scores in the brain tumour group. A potential alternative to overcome this issue would be to use an adaptive staircase procedure for stimulus presentation in the inspection time task, instead of the more commonly employed method of constant stimuli that was used in the present study. The adaptive staircase procedure is an alternative method of presenting the stimuli during the task by which stimulus onset asynchrony (SOA, the amount of time for which the stimulus is presented prior to masking) begins at 320ms when the task commences. A single response error increases SOA by 17ms and three consecutive correct answers are required to decrease SOA by 17ms. The average SOA over eight reversals (i.e. increases or decreases in SOA) of the staircase procedure is taken as the participant's inspection time score (Gregory et al., 2009). This procedure renders the task considerably shorter, with an approximate duration of 10 minutes. As such, this could be a more appropriate version of the inspection time task for use in neuro-oncological patient populations since it could perhaps provide a more accurate assessment method in those patients with reduced concentration and/or extensive visual information processing slowing. Therefore, further research could involve inspection time testing using this adaptive staircase procedure to determine whether visual information processing is slowed in some brain tumour patients to such an extent that the aforementioned validity criterion do not apply.

The present study has a number of strengths compared with similar studies that have previously assessed the presence and extent of cognitive impairment in brain tumour patients. In contrast with many studies that did not recruit a control group for comparison and instead simply compared the brain tumour patient's test scores with normative data in order to classify 'impairment', a particular strength of the present study is that age and sex-matched surgical and healthy control participants were tested on the same battery of tests as the brain tumour patient cohort. This allows us to make a direct comparison of the performance of the three groups who were

administered the same tests, by the same researcher, in precisely the same testing environment, thus eliminating the effects of these potential confounding variables. This methodology also allowed us to both measure and control for each participant's premorbid ability during analysis. Thus, the conclusions made by this and other studies that recruit a matched control group are likely to be more reliable than those made by studies with no control participant group. Moreover, the present study is one of few that detail cognition in brain tumour patients prior to any surgical intervention and as such provides a unique insight into the role of the tumour itself as a cause of cognitive deficits. The majority of studies that have attempted to elucidate the effect of the tumour on cognition in this patient group have recruited patients in the post-operative period. It is therefore difficult to determine the role of the tumour and the role of surgery as a cause of cognitive impairment in these studies (Kayl and Meyers, 2003). Thus this study provides a unique insight into the cognitive function of brain tumour patients at the time of presentation in a large cohort of patients with diverse brain tumours located throughout the brain and suggests that the tumour itself causes cognitive dysfunction in a number of different domains.

A detailed battery of tests assessing several different aspects of cognition, in addition to inspection time as a measure of visual information processing, was administered at baseline testing and this is a further advantage of the study by comparison with a number of similar studies in this area. Some previous studies have relied upon Mini Mental State Examination (MMSE) scores as an indicator of cognitive status (Salander et al., 1995). The MMSE was originally devised as a brief assessment for use with patients with dementia and is not a sensitive measure of cognition in brain tumour patients, many of whom have more subtle deficits in aspects of cognition that are not assessed by the MMSE. As such, many patients with 'normal' MMSE scores may still have cognitive deficits in one or more domains (Meyers and Wefel, 2003). Therefore, the use of a detailed cognitive test battery in the present study provides a more accurate and meaningful assessment of cognitive function than use of a simple, short test like the MMSE that is sensitive to impairments in only the most acutely confused patients.

There also exist some potential weaknesses that are associated with the present study. A considerable proportion of brain tumour patients failed to complete the test battery in its entirety for a number of reasons, particularly by comparison with the healthy control group, the vast majority of whom completed all of the baseline test measures. The reasons for incomplete administration of the test battery primarily included patient fatigue, poor comprehension of task instructions as a result of confusion that prevented task completion, or limited time available for testing due to the competing priorities associated with recruiting patients from the hospital ward. A number of spinal surgery patients also failed to complete the entire battery, and this was usually a result of time constraints since many patients were tested during their pre-admission appointment or on the same day as their operation.

The spinal surgery patients were recruited as a comparison group primarily to control for the effects of surgery and the anxiety associated with an impending operation. Patients admitted for elective spinal surgery were treated within the same department as the brain tumour patients and were therefore easy to recruit and minimised the possible effects of being an inpatient. These are all particular strengths associated with the present study. However, a potential weakness is the failure to recruit a control group of patients with a diagnosis of cancer that does not involve the central nervous system. Klein et al. (2001) propose that all studies of cognition in brain tumour patients should be based upon comparisons with a matched group of other cancer patients to control for the unique stressors and emotions encountered by people with a cancer diagnosis. This methodology would likely have introduced several other confounding variables into the study including variation in the duration of disease and potential treatment effects that may have affected results. Therefore, elective spinal surgery patients were deemed the preferred comparison group in this instance.

The three participant groups, brain tumour, spinal surgery and healthy control, were well matched on the demographic characteristics age, sex and handedness. There were no significant differences in terms of the highest level of education achieved by the members of the brain tumour and spinal surgery groups. However, the healthy

control group had achieved significantly higher levels of education than both of the surgical groups. This could, in part, explain the better performance of the healthy control group on the majority of the baseline measures. However, analysing the data using general linear modelling (analysis of covariance) allowed for inclusion of a measure of premorbid ability to be included as a covariate, thus, the variance in test scores that could be explained by the differing levels of intelligence in the three participant groups was included in each model. Ideally, full-scale IQ scores from the period prior to development of disease would be available for each patient to provide an accurate estimate of premorbid ability from which to gauge the extent of deterioration. However, in the absence of a measure of previous mental ability, the National Adult Reading Test (NART) was used as a measure of premorbid IQ. The NART had been studied extensively and is a proven valid estimate of full-scale IQ (Crawford et al., 1989). Of particular importance for its use in the present study, it has also been successfully validated as a measure of *prior* as opposed to current intellectual ability (Crawford et al., 2001). Thus, this relatively brief, undemanding measure has been shown to provide a valid estimate of premorbid intelligence in a number of patient groups, including patients with Korsakoff's syndrome, Alzheimer's dementia and multi-infarct dementia (Crawford et al., 1988, Bright et al., 2002). McGurn et al. (2004) also validated the use of the NART as an estimate of premorbid IQ in patients with dementia, showing that, after controlling for IQ test score obtained at age 11, mean NART scores in a cohort of the same participants tested at age 80 did not differ significantly, despite the fact that a number of participants had a diagnosis of dementia. Thus, the NART appears to be a reliable estimate of prior ability in patients with mild to moderate dementia. It is therefore reasonable to propose that the NART scores for each participant group in this study are likely to provide an accurate estimate of intelligence, even in the presence of potential mild confusion in some members of the brain tumour cohort. Thus, including NART scores as a covariate in the analyses would likely have effectively controlled for the disparity in levels of intelligence between the healthy and surgical participant groups. Although, as highlighted in Chapter 1.7, no studies have validated the use of the NART in patients with mild dysphasia, few patients in the brain tumour group were judged to have significant speech impairment. We can therefore

assume that mild difficulties with speech production in a very small proportion of the brain tumour patients who participated in the study were unlikely to have significantly affected the overall accuracy of NART scores as a surrogate measure of premorbid function.

There was no significant difference between the two surgical patient groups (brain tumour and spinal surgery) with respect to disability, as measured by the Barthel Disability Index. However, Karnofsky Performance Status scores were paradoxically better for the brain tumour group, despite the superior performance of the spinal surgery control group on the majority of the cognitive tests at baseline. This is likely to be the result of physical impairment and pain due to degenerative spinal disease and replicates the findings of the pilot study in this respect (Zbinden et al., 2006). The majority of the patients in the brain tumour cohort had minimal physical impairment since any focal deficits were likely to have resolved to some extent following dexamethasone (steroid) therapy.

Many medications commonly taken by brain tumour patients, namely dexamethasone and anti-epileptic medications, have been proven to have deleterious effects on cognition, despite their ameliorating effects on focal neurological deficits and seizure frequency/presence, respectively. All of the patients in the brain tumour cohort were taking pre-operative corticosteroids (dexamethasone) and a number of patients were also prescribed anti-epileptic medications as a result of tumour-related seizures. The cognitive and behavioural effects of corticosteroids have been widely studied in a variety of different patient and healthy groups. Corticosteroids have been found to significantly related to behavioural decline in patients with Alzheimer's disease (Aisen et al., 2000) and to be significantly associated with a deterioration in memory and executive functions in groups of healthy male participants (Lupien et al., 1999, Young et al., 1999). Healthy elderly participants with elevated cortisol levels as a result of dexamethasone administration have also been found to exhibit significant cognitive impairments as measured by the Mini Mental State Examination (Kalmijn et al., 1998). It is therefore possible that dexamethasone played some role in contributing to the observed impairment in cognition on the tests administered at

baseline. However, as highlighted by Taphoorn and Klein (2004) in their review, severe cognitive impairment as a result of corticosteroids is rare and it is more likely that dexamethasone works to alleviate cognitive deficits in the majority of brain tumour patients in the same manner by which it improves physical impairment, by reducing intracranial pressure and therefore reducing brain oedema. Thus, the extent of impairment in the brain tumour cohort at baseline may actually have been underestimated, as opposed to exaggerated, due to the potential ameliorating effects of steroid therapy.

Epileptic seizures are a common symptom of brain tumours, particularly in patients with low-grade gliomas (Wessels et al., 2003, Whittle, 2004). Forty-five per cent of the patients in the brain tumour cohort recruited into the present study were thought to have had at last a single seizure at some point prior to surgical intervention and were prescribed anti-epileptic drugs for this reason. Antiepileptic medication has been found to be a risk factor for cognitive impairment, particularly in the domains of information processing speed, psychomotor function, executive function and working memory in a group of long-term low-grade glioma survivors with well controlled seizures (Klein et al., 2003a). The antiepileptic medication prescribed to a high proportion of the brain tumour patients in the present study may therefore have played a role in the observed cognitive impairment at the time of baseline testing. However, given that more than half of the patients in the brain tumour cohort were not taking any form of antiepileptic drug, medication effects alone cannot explain the poorer overall performance in the brain tumour group at the time of baseline testing. Moreover, Tucha et al.(2000) found no difference in the extent of cognitive impairment observed in a group of brain tumour patients who were taking anticonvulsant medication compared with a group of brain tumour patients who were not receiving these drugs. The potential utility of recruiting a control group of patients with epilepsy into the present study in order to allow differentiation of the effects of the tumour from the potential effects of antiepileptic medication was considered. However, recruiting epileptic control patients would have introduced additional confounding variables such as differences in the specific anti-convulsants prescribed, differing dosages between patients and differing lengths of time for

which the medication had been taken and therefore it was decided that recruiting a matched control group of epileptic patients was not feasible.

Mood disorder, such as heightened feelings of anxiety and depression are commonly reported by brain tumour patients and can result in impaired attention and motivation, which in turn may have a deleterious effect on cognitive test performance (Taphoorn and Klein, 2004). The Hospital Anxiety and Depression Scale (HADS) was included as part of the baseline test battery to facilitate comparison of the relative levels of anxiety and depression in each participant group. There were no significant differences between the brain tumour and spinal surgery groups in terms of scores on the anxiety and depression scales of the HADS. This finding is similar to that of Pringle et al. (1999) who found no significant differences in the HADS scores of brain tumour patients and lumbar spinal surgery control patients in the pre-operative period. This would therefore suggest that it is unlikely that the poorer performance of the brain tumour patient group prior to surgery can be attributed to the negative effects of emotional distress on concentration and/or motivation. However, Kilbride et al. (2007) have questioned the sensitivity of the HADS to detect anxiety and depression in neuro-oncological patients. They found that the HADS scores underestimated the presence of anxiety and depression in a group of brain tumour patients compared with the levels of anxiety and depression revealed by content analysis of interviews with the same patients. Thus, the brain tumour group in the present study may actually have experienced more severe emotional distress than their spinal surgery counterparts that was not evident due to the poor sensitivity of the HADS to detect heightened anxiety and depression in neuro-oncological patients. It is for this reason that recruiting a group of patients with a diagnosis of cancer as a control group may have been useful in the present study, in order to provide an adequate control for the emotional effects of such a diagnosis.

The majority of studies that assess cognitive function in brain tumour patients have been unable to elucidate the role of the tumour itself as a cause of cognitive deficits as they involve testing after treatment which may itself be the cause of impairment

(Taphoorn and Klein, 2004). The present study provides a uniquely large data set that provides an insight into the role of the tumour itself in causing cognitive deficits. That the brain tumour group performed significantly worse on tests assessing a variety of different cognitive domains by comparison with surgical and healthy controls supports the findings of one of the few studies that report the prevalence of cognitive impairment prior to any intervention, including surgery, in brain tumour patients. Tucha et al. (2000) assessed a large cohort of patients with tumours located in either the frontal or temporal lobe on a detailed cognitive test battery, prior to any treatment. The authors found that 91% of patients were impaired in at least one area of cognition, with 78% of the cohort exhibiting executive function impairment and 60% showing impairment of memory and attention functions. These authors conclude that most patients with primary or secondary brain tumours present with impairment in some aspect of cognition at the time of admission for, but prior to, neurosurgical intervention. The findings of the present study support the conclusions made by Tucha et al. (2000) and overcome a number of the limitations of the study, namely the failure to recruit an adequate control group for comparison.

6 Post-Operative Function

6.1 Overview of Analysis Procedure

Post-operative (session 2) test performance was compared between the brain tumour, spinal surgery and healthy control groups using general linear modelling (analysis of covariance). Group and sex were entered as fixed effects (between-subjects factors) in the models. Age at the time of testing and NART score were included as covariates for the reasons detailed in Chapter 5.2.1. Baseline inspection time score for each participant was also included as an additional covariate in the models. This effectively allows us to compare any differences between baseline and session 2 performance on each test between the three groups, by controlling for the effects of baseline test performance as well as age, and NART score. There was no significant effect of sex in any of the models for each of the test scores detailed below. Therefore, sex was omitted as a fixed effect in the analyses reported in this chapter. The models were run again with group (brain tumour, spinal surgery and healthy volunteer) specified as the only fixed effect.

In each of the sections that follow, for each test, raw mean baseline and session 2 scores for each group (brain tumour, spinal control, or healthy control) are presented with standard deviations (SD) in brackets. The results of the general linear modelling analysis are then reported. Estimated marginal means (adjusted for age and NART score) and pairwise comparisons are described to highlight the differences between the groups, adjusted for other variables in the models. Least Significant Difference (LSD) tests were used to conduct pairwise post-hoc comparisons.

The mean number of days between baseline testing and session 2 follow-up testing was 6.42 (SD 5.9) for the brain tumour group, compared with 4.39 days (SD 2.3) for the spinal control group and 9.16 days (SD 7.3) for the healthy control group.

Levene's test showed that the group variances were significantly different, $F(2,204) = 11.365$, $p < 0.001$. Therefore, the Kruskal-Wallis test was used to compare differences in the mean number of days between baseline and follow-up in each participant group and this test revealed significant differences between the groups, $H(2) = 33.454$, $p < 0.001$. The difference between the brain tumour and spinal control groups was significant, $U = 1452.0$, $p = 0.001$, suggesting the time to follow-up was significantly longer for the brain tumour group. The difference between the brain tumour and healthy control group was also significant, $U = 1675.0$, $p = 0.001$, suggesting that there was a significantly longer period of time between baseline and session 2 testing in the healthy control group by comparison with the brain tumour group. Comparison of the spinal and healthy control groups was also significant $U = 1218.0$, $p < 0.001$, suggesting that the time to follow-up was significantly longer for the healthy control group than for the spinal control group.

6.2 Results

6.2.1 Inspection Time Scores: All Inspection Time Data

Of the 118 brain tumour patients who participated in the study, there were 60 (50.8%) patients with inspection time scores at both baseline and session 2. Sixty-six (77.6%) of the 85 members of the spinal control group and 76 (95%) of the 80 healthy control group also completed both sessions. All inspection time data, including 'invalid' scores, from participants with scores at both baseline and session 2 follow-up was included in the analysis in the first instance.

The mean total inspection time score at baseline (i.e. total correct /150) for the brain tumour group was 117.6 (SD 20.7) and at session 2 was 113.9 (SD 22.5). The spinal control group scored a mean of 122.2 (SD 13.8) at baseline and a mean of 122.4 (SD 13.4) at session 2. The baseline mean score for the healthy control group was 126.6 (SD 11.3) and at session 2 was 128.5 (SD 12.8). The mean (standard error) scores at baseline and session 2 for each of the three groups are shown in Figure 6.1.

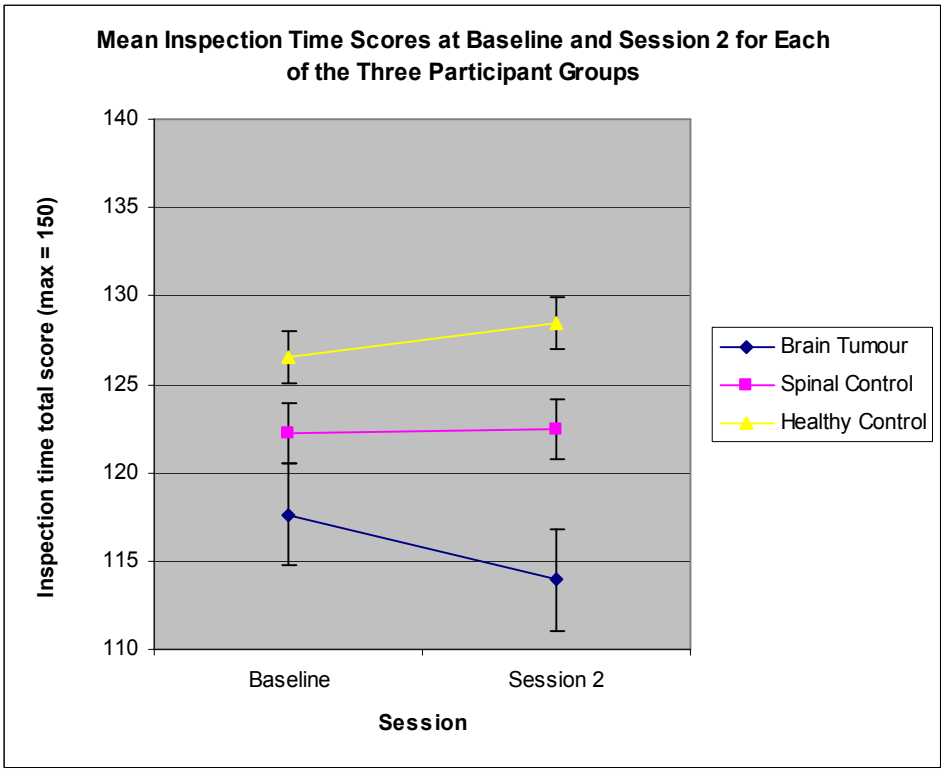


Figure 6.1. Baseline and session 2 inspection time scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with inspection time total score entered as the dependent variable, revealed a significant main effect of the inspection time score at baseline, $F(1,196) = 181.988$, $p < 0.001$, partial $\eta^2 = 0.481$. In the whole sample (i.e. all three participant groups combined), baseline inspection time score was positively correlated with session 2 inspection time score, $r(n = 203) = 0.778$, $p < 0.001$. In the brain tumour group alone, baseline and session 2 inspection time scores were also significantly correlated, $r(n = 60) = 0.709$, $p < 0.001$. Similar correlations were obtained for the spinal control group, $r(n = 67) = 0.813$, $p < 0.001$ and for the healthy control group, $r(n = 76) = 0.855$, $p < 0.001$. This therefore shows that those participants with better baseline inspection time scores also performed significantly better at session 2 follow-up. There was a significant main effect of the covariate age on session 2 inspection time scores, $F(1,196) = 9.045$, $p = 0.003$, partial $\eta^2 = 0.044$. Older participants tended to have lower inspection time scores than younger participants. There was a significant main effect of the covariate NART score in the model, $F(1,196) = 4.629$, $p = 0.033$, partial $\eta^2 = 0.023$. Participants with higher (i.e. poorer) NART scores had lower session 2 inspection time scores. Participant group

had a significant main effect in the model that included the effects of the covariates age, NART and baseline inspection time score, $F(2,196) = 6.241$, $p = 0.002$, partial $\eta^2 = 0.060$. That is, in the presence of the effects of age, NART score and baseline inspection time score, there was a significant difference between the performance of the three participant groups on inspection time at post-operative (session 2) testing. These group differences, as planned, were explored further.

The estimated marginal mean session 2 inspection time score, adjusted for age, NART and baseline inspection time score—derived from the above-described general linear model—was 118.1 (SE 1.4) for the brain tumour group, 123.1 (SD 1.3) for the spinal control group and 124.6 (SE 1.3) for the healthy control group. Pairwise comparisons using Least Significant Difference tests show that the brain tumour group performed significantly worse than the spinal control group ($p = 0.008$) at session 2. The brain tumour group were also significantly worse the healthy control group ($p = 0.001$). There was no significant difference between the spinal and healthy control group performance on session 2 inspection time scores ($p = 0.424$).

Therefore, there was a significantly greater relative deterioration in inspection time performance in the brain tumour group between baseline and session 2 testing by comparison with both the spinal surgery and healthy control groups.

6.2.2 Inspection Time Scores: Valid Inspection Time Data

There were 52 (44.1%) members of the total cohort of 118 brain tumour patients with 'valid' baseline inspection time scores, who also had follow-up inspection time (valid or invalid) scores at session 2. 'Valid' inspection time scores are data from those participants who scored at least 17/20 on the longest two durations, as detailed in chapter 2.2. Data from patients who had valid scores at baseline but invalid scores at session 2 were also included here as these participants were deemed to have proven their comprehension of the task on the first occasion. Sixty-two (72.9%) of

the 85 spinal controls and 74 (92.5%) of the total cohort of 80 healthy controls also met these validity criteria.

The mean baseline inspection time score for the brain tumour group, when only valid scores were included, was 123.8 (SD 12.7), and following surgery at session 2 was 117.5 (SD 19.8). For the spinal group, the mean baseline score was 124.7 (SD 10.1) and at session 2 was 124.5 (SD 10.6). The healthy control group scored a mean of 127.5 (SD 10.0) at baseline and a mean of 129.6 (SD 10.8) at session 2. The mean (standard error) scores at baseline and session 2 for each of the groups are shown in Figure 6.2.

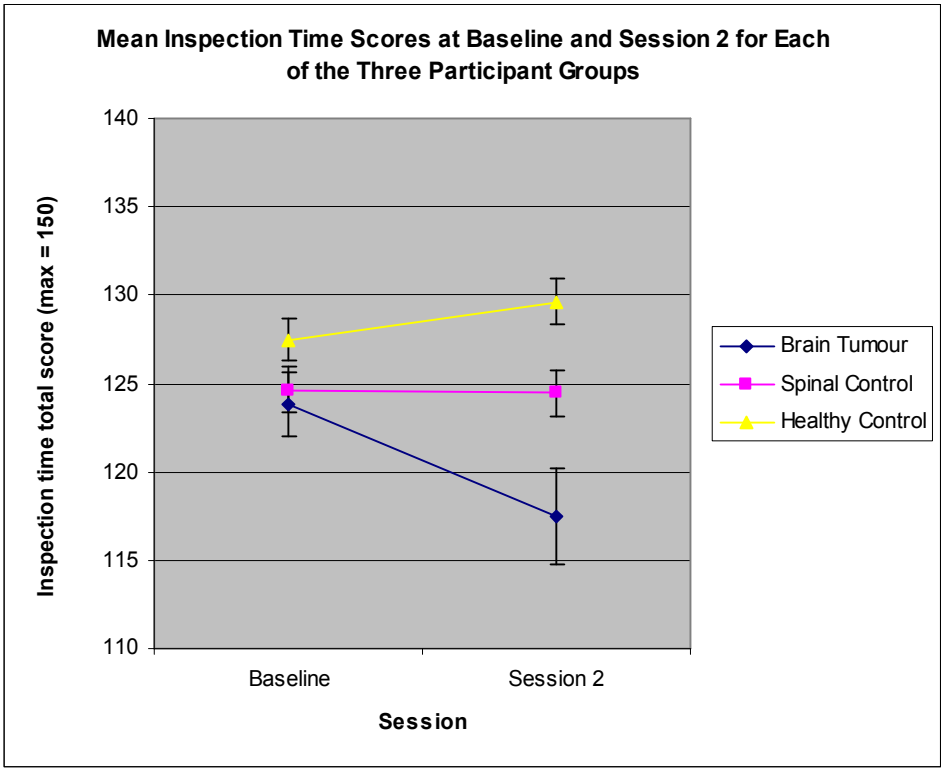


Figure 6.2. Baseline and session 2 inspection time scores for the three participant groups including only data from participants with valid scores at baseline. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with total inspection time score at session 2 entered as the dependent variable, revealed a significant main effect of the baseline inspection time score, $F(1,182) = 90.587$, $p < 0.001$, partial $\eta^2 = 0.332$. In the whole sample combined, valid baseline inspection time scores were correlated with session 2 inspection time scores, $r(n = 188) = 0.675$, $p < 0.001$. In the brain tumour group

alone, valid baseline score was correlated with session 2 score, $r(n = 52) = 0.611$, $p < 0.001$. This was also the case for the spinal control group, $r(n=62) = 0.694$, $p < 0.001$, and for the healthy control group, $r(n = 74) = 0.805$, $p < 0.001$. Those participants who scored better at baseline performed significantly better at session 2 follow-up. There was a significant main effect of the covariate age on session 2 inspection time scores, $F(1,182) = 9.203$, $p = 0.003$, partial $\eta^2 = 0.048$. Older participants tended to have poorer scores than younger participants. The covariate NART score also had a significant main effect in the model, $F(1,182) = 6.606$, $p = 0.011$, partial $\eta^2 = 0.035$. Poorer performance on the NART was significantly related to lower inspection time scores at session 2. Participant group had a significant main effect in the model that included the effects of age, NART score and baseline inspection time scores, $F(2,182) = 9.521$, $p < 0.001$, partial $\eta^2 = 0.095$. That is, in the presence of the effects of age, NART score and baseline inspection time score for those participants with 'valid' test scores; there was a significant difference between the performance of the three participant groups on inspection time testing at session 2. These between-groups differences were explored further.

The estimated marginal mean session 2 inspection time scores for participants with valid scores at baseline, adjusted for age, NART and baseline score and derived from the general linear model described above were 119.3 (SD 1.4) for the brain tumour group, 125.7 (SD 1.3) for the spinal control group, and 127.3 (SD 1.2) for the healthy control group. Pairwise comparisons, using Least Significant Difference tests, revealed that the brain tumour group performed significantly worse than the spinal control group ($p = 0.001$) and significantly worse than the healthy control group ($p < 0.001$). There was no significant difference between the performance of the spinal and healthy control groups ($p = 0.406$).

Therefore, in the cohort of patients with valid inspection time scores at baseline, there was a significantly greater relative deterioration in the brain tumour group between baseline and session 2 testing by comparison with both control groups (spinal surgery and healthy volunteer).

6.2.3 Digit Symbol Coding

Of the 118 brain tumour patients who participated in the study, 60 (50.8%) completed the digit symbol-coding task at both baseline and post-operatively at session 2. Sixty-five of the 85 spinal control patients (76.5%) and 75 of the 80 healthy volunteer controls (93.8%) completed the task at both baseline and session 2. Only data from participants with test scores at both sessions were included in the analyses.

At baseline, the brain tumour patients scored a mean of 62.2 (SD 21.3), compared with a mean of 58.2 (SD 21.5) at session 2, post-operatively. The spinal group scored a mean of 69.2 (SD 17.3) at baseline compared with a mean of 66.9 (SD 20.0) at post-operative follow-up (session 2). The healthy control group had a mean baseline score of 76.6 (SD 16.0) and a mean session 2 score of 81.9 (SD 18.2). The mean (standard error) scores at baseline and session 2 for each of the groups are shown in Figure 6.3.

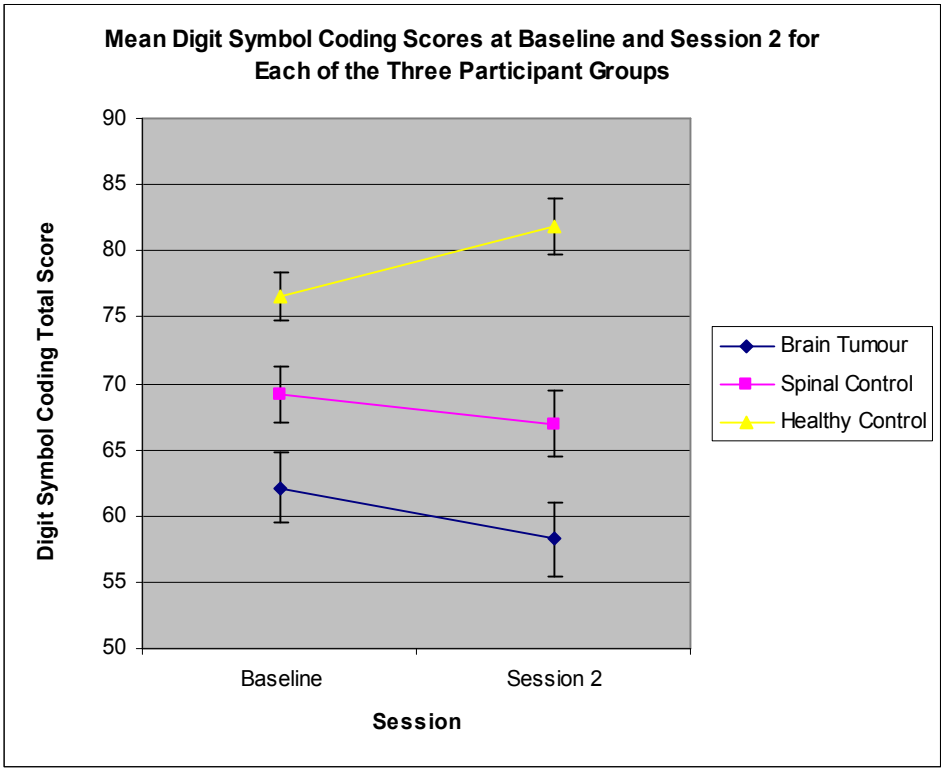


Figure 6.3. Baseline and session 2 digit symbol coding scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with digit symbol coding score at session 2 as the dependent variable, revealed a significant main effect of baseline digit symbol coding score, $F(1,194) = 330.211$, $p < 0.001$, partial $\eta^2 = 0.630$. In the whole sample combined, digit symbol coding score at baseline was correlated with digit symbol coding score at session 2, $r(n = 201) = 0.862$, $p < 0.001$. The correlation between baseline and session 2 digit symbol coding scores for the brain tumour group alone was also significant, $r(n = 60) = 0.892$, $p < 0.001$. Similar significant correlations were obtained for the spinal control group, $r(n = 66) = 0.726$, $p < 0.001$, and also for the healthy control group, $r(n = 75) = 0.942$, $p < 0.001$. Those participants with higher scores at baseline on digit symbol coding performed significantly better on the same test at session 2 follow-up than those participants with lower baseline scores. There was no significant main effect of the covariate age on digit symbol scores at session 2, $F(1,194) = 1.653$, $p = 0.200$, partial $\eta^2 = 0.008$. Similarly, NART score had no significant main effect on session 2 scores on the digit symbol coding measure, $F(1,194) = 0.990$, $p = 0.321$, partial $\eta^2 = 0.005$. Participant group had a significant main effect on digit symbol coding scores at session 2, in the model that included the effects of baseline test scores, age and NART score, $F(2,194) = 13.662$, $p < 0.001$, partial $\eta^2 = 0.123$. This suggests that, in the presence of the effects of baseline test scores, age and NART score, there was a significant difference between the performance of the three participant groups on digit symbol coding at session 2 follow-up. These differences were explored further, using pairwise comparisons as planned.

The estimated marginal mean session 2 digit symbol coding scores, adjusted for the effects of age, NART score and baseline digit symbol coding score – derived from the above-described general linear model - were 65.4 (SE 1.4) for the brain tumour group, 67.7 (SE 1.3) for the spinal surgery control group and 75.5 (SE 1.3) for the healthy control group. Pairwise comparisons using Least Significant Difference tests demonstrate that the brain tumour group performed significantly worse than the healthy control group at session 2 ($p < 0.001$). However, there was no significant difference between the brain tumour group and the spinal control group ($p = 0.231$). The spinal control group performed significantly worse than the healthy control

group at session 2 ($p < 0.001$). This shows that the brain tumour and spinal groups both deteriorated significantly more on digit symbol coding at session 2 compared than the healthy control group, who had a higher session 2 score compared with baseline performance.

6.2.4 Nine Hole Peg Test (Right Hand, EFIT)

In the brain tumour group, 62 of 118 patients (52.5%) completed the Nine Hole Peg Test (NHPT) at both baseline and at session 2 post-operatively. Sixty-four of the total cohort of 85 spinal control patients (75.3%) and 76 of the total cohort of 80 healthy volunteer controls (95.0%) completed the test on both occasions.
The brain tumour group scored a mean of 13.8 seconds (SD 2.9) at baseline on this test, and a mean of 14.5 seconds (SD 2.8) at post-operative follow-up (session 2). The corresponding mean score for the spinal group at baseline was 12.9 seconds (SD 2.9) and was 13.5 seconds (SD 2.7) at session 2. The healthy control group scored a mean of 12.3 seconds (SD 1.5) at baseline and 12.0 seconds (SD 1.7) at session 2. The mean (standard error) scores at baseline and session 2 for each of the three groups are shown in Figure 6.4.

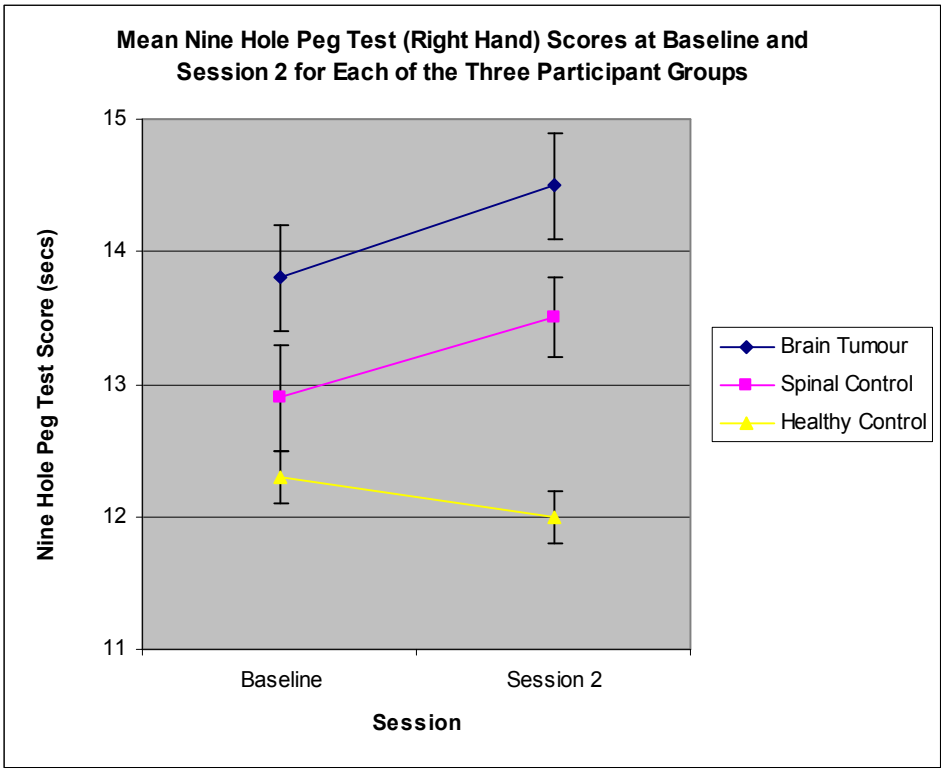


Figure 6.4. Baseline and session 2 nine hole peg test (right hand) scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with right-hand NHPT score at session 2 entered as the dependent variable, revealed a significant main effect of baseline score on the right-hand NHPT, $F(1,196) = 161.119$, $p < 0.001$, partial $\eta^2 = 0.451$. In the whole sample of participants recruited into the study, baseline nine hole peg test score (right hand) was correlated with session 2 NHPT score (right hand), $r(n = 203) = 0.750$, $p < 0.001$. In the brain tumour sample alone, there was a significant correlation between the baseline and session 2 scores, $r(n = 62) = 0.741$, $p < 0.001$. The correlations were also significant for the spinal control group, $r(n = 65) = 0.802$, $p < 0.001$ and the healthy control group, $r(n = 76) = 0.578$, $p < 0.001$. Therefore those participants who took longer to complete the task at baseline were significantly more likely to take longer to complete the task at session 2. There was a significant main effect of the covariate age on session 2 NHPT (right hand) scores, $F(1,196) = 7.110$, $p = 0.008$, partial $\eta^2 = 0.035$. Older participants tended to take longer to complete the task than younger participants. There was no significant main effect of the covariate NART score on session 2 NHPT (right hand) scores, $F(1,196) = 0.191$, $p = 0.662$, partial $\eta^2 = 0.001$. Participant group had a significant main effect in the model that included

the effects of the covariates age, NART and baseline score on the NHPT, $F(2,196) = 13.156$, $p < 0.001$, partial $\eta^2 = 0.118$. Therefore, in the presence of the effects of age, NART and baseline right-hand NHPT score, there was a significant difference between the performance of the three participant groups on right-hand NHPT at session 2 follow-up. These differences were explored further, using post-hoc tests as planned.

The estimated marginal mean session 2 right-hand NHPT scores, adjusted for age, NART and baseline score on the same test – derived form the above-described general linear model – were 13.9 (SE 0.2) for the brain tumour group, for the spinal control group 13.5 (SE 0.2) and 12.4 (SE 0.2) for the healthy control group. Pairwise comparisons using Least Significant Difference tests show that the brain tumour did not perform significantly differently from the spinal control group at session 2 ($p = 0.199$). The brain tumour group did however perform significantly worse than the healthy control group ($p < 0.001$). The spinal control group also performed significantly worse at session 2 than the healthy control group ($p < 0.001$).

Therefore, the brain tumour group and spinal control group showed a greater relative deterioration in nine hole peg test (right hand) performance between baseline and session 2 than the healthy control group. There was no significant difference in the extent of deterioration between the brain tumour and spinal surgery groups.

6.2.5 Nine Hole Peg Test (Left Hand, EFIT)

Sixty-one patients in the brain tumour group completed this test at both baseline and session 2 follow-up. A further 65 spinal controls and 76 healthy controls also completed the NHPT with the left hand on both occasions.

The mean time taken to complete this test at baseline was 15.3 seconds (SD 3.6) for the brain tumour group and at post-operative follow-up (session 2) the mean time taken was 18.5 seconds (SD 21.1). The spinal control group had a mean of 14.2 seconds (SD 3.3) at baseline and a mean score of 15.3 seconds (SD 5.5) at session 2.

At baseline, the healthy control group scored a mean of 12.9 seconds (SD 2.0), with a mean of 12.6 seconds (SD 1.9) at session 2. The mean (standard error) scores at baseline and session 2 for each of the three groups are shown in Figure 6.5.

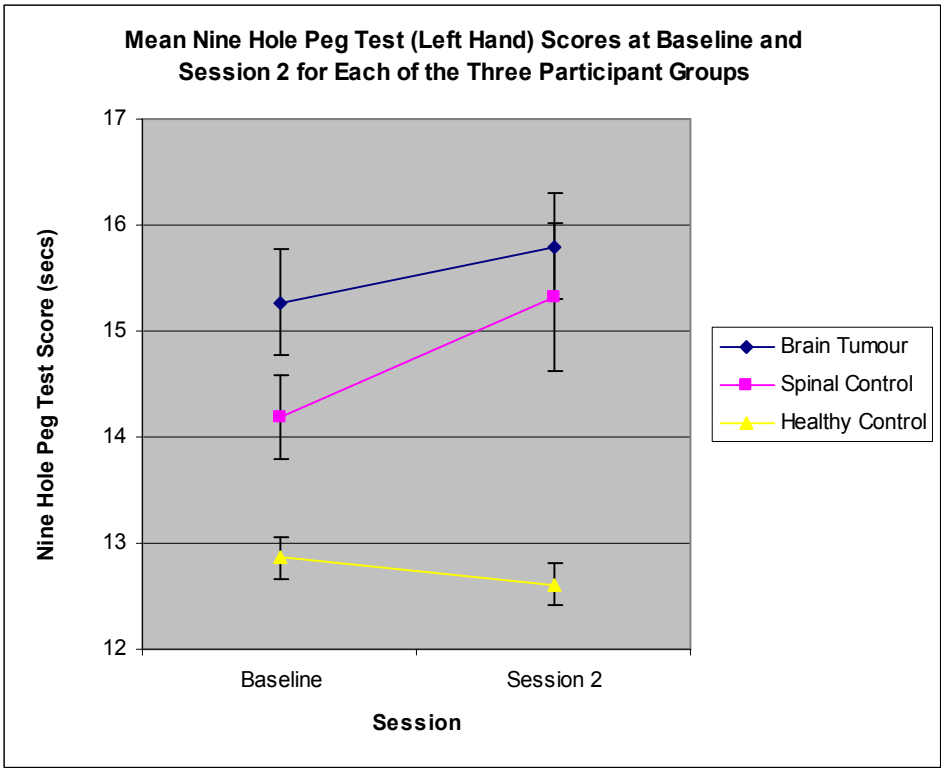


Figure 6.5. Baseline and session 2 nine hole peg test (left hand) scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with left hand NHPT score entered as the dependent variable, revealed a significant main effect of the same test score at baseline, $F(1,196) = 103.515$, $p < 0.001$, partial $\eta^2 = 0.346$. In the whole sample (i.e. all three participant groups combined), baseline left hand NHPT score was positively correlated with session 2 score on the same measure, $r(n = 203) = 0.673$, $p < 0.001$. In the brain tumour group alone, baseline and session 2 score on the left hand nine hole peg test were also positively correlated, $r(n = 61) = 0.859$, $p < 0.001$. Similar significant correlations were obtained for the spinal control group, $r(n = 66) = 0.529$, $p < 0.001$, and for the healthy control group, $r(n = 76) = 0.683$, $p < 0.001$. This therefore shows that those participants with faster baseline scores on this test were more likely to also complete the test faster at session 2. There was no significant main effect of age on session 2 left hand NHPT scores, $F(1,196) = 1.114$. $p = 0.293$,

partial $\eta^2 = 0.006$. There was also no significant main effect of NART score on session 2 scores on this test, $F(1,196) = 2.027$, $p = 0.156$, partial $\eta^2 = 0.010$. Participant group had a significant main effect in the model that included the effects of the covariates age, NART and baseline test score on the measure, $F(2,196) = 6.977$, $p = 0.001$, partial $\eta^2 = 0.066$. In the presence of the effects of age, NART score and baseline left hand nine hole peg test score, there was a significant difference between the performance of the three participant groups on this measure at session 2. The between-groups differences were examined further, using pairwise comparisons, as planned.

The estimated marginal mean left-hand NHPT scores at session 2, adjusted for age, NART and baseline score on the same test and derived from the above-described general linear model were 14.9 (SE 0.4) for the brain tumour group, 15.3 (SE 0.4) for the spinal control group and 13.3 (SE 0.4) for the healthy control group. Pairwise comparisons using Least Significant Difference tests showed there to be no significant difference between the performance of the brain tumour and spinal control groups on this test at session 2 ($p = 0.402$). The brain tumour group were significantly worse than the healthy control group ($p = 0.007$) and the spinal control group were also significantly worse than the healthy control group ($p < 0.001$). Therefore, there was a significantly greater relative deterioration on the left hand NHPT in the brain tumour and spinal control group between baseline and session 2 testing by comparison with the healthy control group but no significant difference between the brain tumour and spinal surgery groups.

6.2.6 Williams Delayed Recall Test (EFIT)

This test was completed at both baseline and session 2 by 62 of the 118 members of the brain tumour group (52.5%), 65 of the 85 spinal controls (76.5%) and 76 of the 80 healthy controls (95.0%).

The mean scores for the brain tumour group who completed both testing sessions were 8.7 (SD 5.2) at baseline and 11.4 (SD 6.1) at session 2. The spinal group scored

a mean of 5.1 (SD 3.7) at baseline, and 7.3 (SD 4.0) at session 2. The mean baseline score for the healthy control group was 4.5 (SD 3.6), compared with 4.1 (SD 3.1) at session 2 follow-up. The mean (standard error) scores at baseline and session 2 for each of the three groups are shown in Figure 6.6.

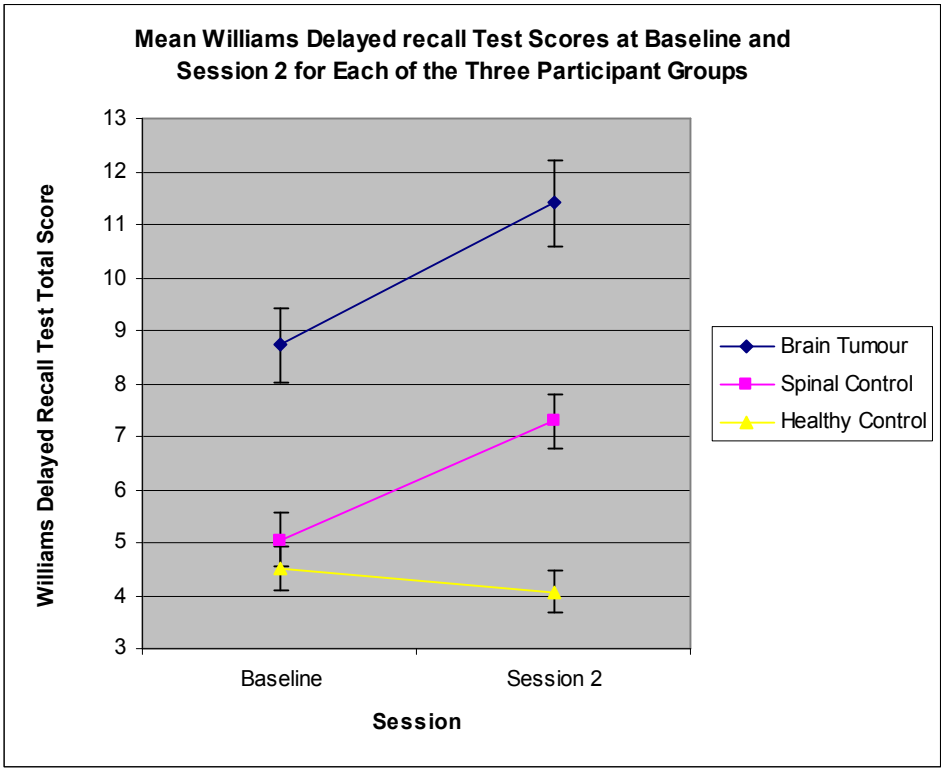


Figure 6.6. Baseline and session 2 Williams Delayed Recall Test scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with Williams Delayed Recall Test (WDRT) score entered as the dependent variable, revealed a significant main effect of the baseline WDRT, $F(1,197) = 50.304$, $p < 0.001$, partial $\eta^2 = 0.203$. In the whole sample of all three participant groups combined, baseline WDRT score was positively correlated with the same test score at session 2, $r(n = 204) = 0.612$, $p < 0.001$. In the brain tumour group alone, baseline and session two scores on this measure were also positively correlated, $r(n = 62) = 0.495$, $p < 0.001$. The correlation between the two scores for the spinal group was $r(n = 66) = 0.544$, $p < 0.001$; and for the healthy control group was $r(n = 76) = 0.601$, $p < 0.001$. Those participants who had lower scores at baseline on the WDRT also tended to have lower scores on the same measure at session 2. There was a significant main effect of the covariate age on session 2 scores

on the test, $F(1,197) = 4.521$, $p = 0.035$, partial $\eta^2 = 0.022$. Older participants tended to perform less well than younger participants. There was also a significant main effect of NART score in the model, $F(1,197) = 8.565$, $p = 0.004$, partial $\eta^2 = 0.042$. Participants with higher (poorer) NART scores tended to have higher (poorer) scores on the WDRT. Participant group had a significant main effect in the model that included the effects of the covariates age, NART score and baseline WDRT score, $F(2,197) = 19.102$, $p < 0.001$, partial $\eta^2 = 0.162$. This suggests that, in the presence of the effects of age, NART score and baseline score on the WDRT, there was a significant difference between the performances of the three groups at session 2. As planned, these group differences were explored using post-hoc pairwise comparisons (Least Significant Difference tests).

The estimated marginal mean session 2 scores on the WDRT, adjusted for age, NART and baseline score on the same test and derived from the above-described general linear model, were 9.7 (SE 0.5) for the brain tumour group, 7.5 (SE 0.5) for the spinal control group and 5.2 (SE 0.5) for the healthy control group. Pairwise comparisons using Least Significant Difference tests show that the brain tumour group performed significantly worse than the spinal control group at session 2 ($p = 0.002$). The brain tumour group were also significantly worse than the healthy control group ($p < 0.001$). The spinal group were significantly worse than the healthy control group ($p = 0.001$). This suggests there was a significantly greater relative deterioration in WDRT performance in the brain tumour group between baseline and session 2 by comparison with both the spinal surgery and healthy control groups. The spinal group also deteriorated significantly more by comparison with the healthy control group.

6.2.7 Timed Ten Metre Walk (EFIT)

This test was completed at both baseline and session 2 by 56 of the 118 members of the brain tumour cohort (n = 118, 47.5%), 57 of the 85 spinal control patients (67.1%) and 72 of the 80 healthy volunteer controls (90%).

The brain tumour patients who completed this test on both occasions (baseline and session 2) scored a mean of 6.7 seconds (SD 1.4) at baseline and a mean of 7.1 seconds (SD 1.7). For the spinal group, mean baseline score was 7.8 seconds (SD 2.1) and at session 2 was 8.4 seconds (SD 2.6). The healthy control group mean at baseline was 6.0 seconds (SD 1.2) and was 6.0 seconds (1.0) at session 2. The mean (standard error) scores at baseline and session 2 for each of the three groups are shown in Figure 6.7.

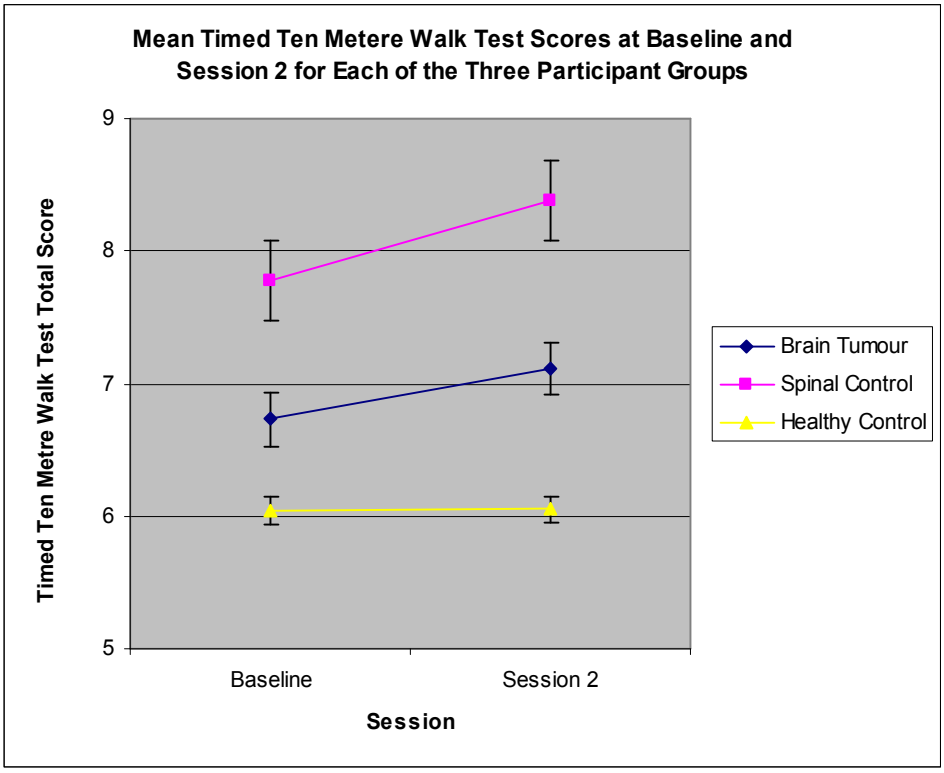


Figure 6.7. Baseline and session 2 Timed Ten Metre Walk test scores for the three participant groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with session 2 timed ten metre walk score entered as the dependent variable, revealed a significant main effect of the baseline score on this measure, $F(1,179) = 89.695$, $p < 0.001$, partial $\eta^2 = 0.334$. In the whole sample (i.e. all three participant groups combined), baseline ten metre walk score was positively correlated with session 2 ten metre walk score, $r(n = 186) = 0.695$, $p < 0.001$. In the brain tumour group alone, baseline and session 2 scores on this test were also positively correlated, $r(n = 56) = 0.731$, $p < 0.001$. This was also the case for the spinal control group, $r(n = 58) = 0.574$, $p < 0.001$; and also for the healthy control group, $r(n = 72) = 0.660$, $p < 0.001$. Those patients who had better (i.e. quicker) scores on this measure at baseline also performed significantly better at follow-up. There was no significant main effect of either of the covariates age, $F(1,179) = 3.112$, $p = 0.079$, partial $\eta^2 = 0.017$; or NART score in the model, $F(1,179) = 0.136$, $p = 0.713$, partial $\eta^2 = 0.001$. Participant group had a significant main effect in the model that included the effects of baseline ten metre walk score, NART and age, $F(2,179) = 8.257$, $p < 0.001$, partial $\eta^2 = 0.084$. In the presence of the effects of age, NART score and baseline ten metre walk score, there was a significant difference between the three participant groups on the timed ten metre walk at session 2. To explore these differences between groups, pairwise comparisons were used as planned.

The estimated marginal mean session 2 timed ten metre walk scores, adjusted for age, NART and baseline ten metre walk score, and derived from the general linear model described above, were 7.1 (SE 0.2) for the brain tumour group, 7.7 (SE 0.2) for the spinal control group and 6.5 (SE 0.2) for the healthy control group. Pairwise comparisons using Least Significant Difference tests show that the spinal group performed significantly worse than the brain tumour group ($p = 0.031$). The spinal group were also significantly worse than the healthy control group ($p < 0.001$). The brain tumour group were significantly worse than the healthy control group ($p = 0.027$) on this measure at session 2. Therefore, the spinal control group showed significantly greater deterioration on the timed ten metre walk between baseline and session 2 by comparison with both the brain tumour group and the healthy control

group. The brain tumour group also deteriorated significantly more between baseline and session 2 than the healthy control group.

Table 6.1 Estimated marginal mean session 2 test scores, adjusted for age, NART score and baseline test score, and post hoc pairwise comparisons.

Test	Estimated marginal mean score (SE) at session 2*			Pairwise comparisons – p value for significance		
	Brain Tumour Group	Spinal Control Group	Healthy Control Group	Brain Tumour vs. Spinal Control	Brain Tumour vs. Healthy Control	Spinal Control vs. Healthy Control
Inspection Time (All Data)[1]	118.1 (1.4)	123.1 (1.3)	124.6 (1.3)	0.008	0.001	0.424
Inspection Time (Valid Data)[1]	119.3 (1.4)	125.7 (1.3)	127.3 (1.2)	0.001	< 0.001	0.406
Digit Symbol Coding[1]	65.4 (1.4)	67.7 (1.3)	75.5 (1.3)	0.231	< 0.001	< 0.001
Nine Hole Peg Test – Right Hand (EFIT)[2]	13.9 (0.2)	13.5 (0.2)	12.4 (0.2)	0.199	< 0.001	< 0.001
Nine Hole Peg Test – Left Hand (EFIT)[2]	14.9 (0.4)	15.3 (0.4)	13.3 (0.4)	0.402	0.007	< 0.001
Williams Delayed Recall Test (EFIT)[2]	9.7 (0.5)	7.5 (0.5)	5.2 (0.5)	0.002	< 0.001	0.001
Timed Ten Metre Walk (EFIT)[2]	7.1 (0.2)	7.7 (0.2)	6.5 (0.2)	0.031	0.027	< 0.001

*Estimated marginal mean score, adjusted for age, NART score and test score at baseline. SE = Standard Error.
[1] Higher scores represent better function. [2] Higher scores represent worse function.

6.3 Discussion

Overall, the brain tumour group were found to deteriorate significantly more between baseline and session 2 on the inspection time task than both the spinal surgery control group and the healthy control group. Neither the spinal surgery group nor the healthy control group deteriorated significantly on inspection time between baseline and session 2. This was the case when all inspection time data were included in the analysis, and the difference between the brain tumour group and the two control groups was even larger when only 'valid' inspection time scores were included in the analysis. The effect size on both occasions was moderate. This suggests that the inspection time task a sensitive measure of the effects of brain tumour surgery on visual information processing.

On all of the other tests that comprised the post-operative test battery, (i.e. digit symbol coding and the Edinburgh Functional Impairment Tests, EFIT), both the brain tumour and spinal surgery groups deteriorated significantly when tested post-operatively by comparison with the healthy control group who were tested on a second occasion. This suggests that surgery in general had a deleterious effect on patient function, as measured by these tests. The effect size was moderate – large for each of these tests. This suggests that the inspection time task is not unique in its ability to detect the effects of surgery in brain tumour patients. However, compared with the spinal surgery control group, the brain tumour group deteriorated significantly more on only the Williams Delayed Recall Test (WDRT) from the EFITs. Thus, inspection time and the WDRT appeared to be particularly sensitive to the effects of brain tumour surgery, over and above the general deleterious effects of surgery detected by the other measures.

There were a number of advantages associated with enrolling a group of patients who were admitted for elective spinal surgery as a control group into the study. Namely, the spinal surgery patients provide a control group for general surgery-related variables such as anaesthesia and fatigue. A general post-operative cognitive dysfunction has been reported in a number of different surgical patient groups, and in

particularly in elderly surgical patients (Bekker and Weeks, 2003, Rasmussen, 1999, Moller, 1997). Thus, without a matched surgical group for comparison, any post-operative deterioration in the brain tumour group could simply be explained as the result of the general post-operative cognitive dysfunction that is common to many different surgical populations. However, since the spinal control group did not show the same deterioration in inspection time and WDRT performance post-operatively, the post-surgical deterioration following brain tumour surgery on these measures cannot simply be attributed solely to the effects of a general post-operative cognitive dysfunction. It is more likely that the specific effect of tumour removal or biopsy caused a general slowing in processing speed, as measured by the inspection time test and also impairment in delayed recall as measured by the WDRT. Certainly, previous studies examining the post-operative cognitive function of brain tumour patients have shown a post-operative worsening of psychomotor function in low-grade glioma patients (Reijneveld et al., 2001) and a transient general cognitive impairment was observed by Teixidor et al. (2006). However, other studies have also found there to be no significant post-operative change in cognition, for example, in patients with a meningioma (Tucha et al., 2003). In a study with a methodology that broadly reflects that of the present study, Habets et al. (2008) found surgical resection to have no negative effect on cognition in a heterogeneous sample of brain tumour patients, unless there were specific post-operative complications. Indeed, in direct contrast with the findings of our study, these authors observed an improvement in mental speed and perceptual abilities in the post-operative period. However, the study was limited by a very small sample size (n = 21) and considerable variation in the time at which post-operative assessment was carried out (4 – 104 days).

An additional strength associated with the methodology of the present study is that the tests that comprised the battery administered at session 2 were chosen specifically because they have either been shown to have minimal practice effects or have alternate forms that make the task particularly suitable for repeated assessment. However, practice effects may exist to some extent in the tests administered at session 2. For example, an improvement in inspection time performance over repeated testing has previously been demonstrated in healthy volunteers (Bors et al.,

1999b). Therefore, a further strength of the study design was the inclusion of a group of non-surgical, healthy volunteer control participants to allow examination of the pure effect of practice which is of importance when interpreting cognitive test scores following repeated administration of a task. The healthy control group performed marginally better at session 2 compared with baseline performance on a number of the tests, namely inspection time and digit symbol coding, confirming previous studies that have observed practice effects associated with repeated administration of the tests (Bors et al., 1999b, Beres and Baron, 1981, Bors et al., 1999a). Thus, even in the presence of the potential positive effects of practice on test performance, the surgical cohorts deteriorated on a number of the tests at session 2. Overall, the time between baseline testing and post-operative (session 2) follow-up was longer for the brain tumour group compared with the spinal control group and was longest for the healthy control group. Studies have shown that the positive effects of practice on test performance are most likely to be seen between the first and second administration of a cognitive measure (Benedict and Zgaljardic, 1998, Collie et al., 2003). In fact Collie et al. (2003) identify the potential utility of conducting two baseline testing sessions, and subsequently excluding of the results from the first baseline session in order to minimise the observed practice effects. This methodology would not, however, have been feasible in the present study since recruiting patients for even a single baseline testing session prior to surgery was problematic due to the time constraints associated with recruiting inpatients from a busy hospital ward. The results of these aforementioned studies lead us to expect that the participant group with the shortest elapsed time between baseline and session 2 testing would be most likely to exhibit practice effects. However, since the healthy control group had the longest duration between the two sessions, yet exhibited the greatest level of improvement at session 2, we can conclude that practice effects had no significant impact on the conclusions reached in this study. To control for within group variation in terms of the interval between baseline and session 2 testing, one study of pulmonary bypass patients employed very strict follow-up criteria whereby only the results from patients who completed the second testing session exactly 8 days post-operatively, as scheduled, were included in the analysis (Pugsley et al., 1994). Given the variation in terms of recovery and discharge times for the surgical patients in this

study, such a method would not have been feasible and would have lead to the exclusion of a very high proportion of patients, many of whom may have experienced significant impairment post-operatively. Moreover, this methodology is not necessary in the present study given that practice effects were unlikely to have had any significant effect on the conclusions made in the study.

The prospective nature of the present study gives the present study an advantage over others that report the presence of post-operative impairment in brain tumour patients since cognitive test scores were available from the pre-operative period to allow a direct comparison with post-operative test scores from the same patients. Thus, a direct assessment of the effects of surgery on inspection time, and other, function can be made by comparison with the same patient's test scores in the pre-operative period. Several previous studies attempting to elucidate the potential role of surgery in causing cognitive deficits have attempted to do so by comparing the test scores of patients who have already had surgery with either those of age and sex-matched patients who are yet to have surgical intervention (Reijneveld et al., 2001) or with age and sex-matched healthy control participants (Klein et al., 2001). However, several confounding variables, such as differences in premorbid intelligence and differences in terms of tumour size and location limit the conclusions that can be drawn from the aforementioned studies.

It has been suggested that studies of post-operative cognitive dysfunction should aim to minimise variability between examiners and within individual participants as much as possible (Rasmussen et al., 2001). This study was designed to meet these aims as much as was feasibly possible. Firstly, all tests were administered at each session and to each group by the same researcher thus minimising potential examiner effects and eliminating the need to assess inter-rater reliability. Each testing session took place in a controlled testing environment that aimed to minimise external distractions, under constant artificial lighting conditions. However, there was considerable variability in terms of the time of day during which testing took place. Baseline testing often took place later in the day following admission to the hospital ward, whereas post-operative testing generally took place in the morning on the day

of discharge. Additionally, it was impossible to avoid the potential effects of increased fatigue that the surgical cohorts likely experienced post-operatively, following an inpatient stay on a hospital ward. However, given that both the brain tumour and spinal surgery groups were admitted to the same neurosurgical wards, it is likely that the extent of fatigue was similar between the two groups, although no formal measure of this variable was made. Therefore, the greater relative post-operative deterioration on inspection time and the WDRT cannot be attributed to the effects of fatigue or to the variation in terms of baseline and session 2 testing times.

Despite the advantages associated with recruiting both the spinal surgery and healthy volunteer control groups into the study, comparing the scores of the brain tumour cohort with two control groups raises specific issues associated with data analysis. To compare the baseline and session 2 follow-up scores for each participant group general linear modelling (analysis of covariance) was used since this allowed us to control for important variables that have previously been shown to have a significant effect on cognitive test performance. These variables include National Adult Reading Test score as a measure of premorbid intelligence and age at the time of testing since both variables have been found to correlate with inspection time (Deary and Stough, 1996) and with cognitive test performance in general (Lezak et al., 2004). Least Significant Difference (LSD) pairwise comparisons were specified during analysis in order to identify where any specific between-groups differences lie, should a significant overall effect of group be found in the model. Running LSD comparisons is equivalent to running multiple t-tests between each pair of groups, since there is no adjustment made to account for multiple comparisons being made. Thus, the probability of obtaining a significant result is increased. However, given that there were only three participant groups, and therefore three between-group comparisons to run, LSD comparisons were deemed the preferred method of analysis in this instance. A potential alternative would have been to use the more conservative Bonferroni correction which would have increased the probability of Type II error. Given that, based upon the results of a pilot study (Zbinden et al., 2006), it was hypothesised that the group of brain tumour patients would show significantly

greater post-operative deterioration, LSD comparisons were the most appropriate post hoc comparisons to use in this instance.

As highlighted in Chapter 1.5.3, some medications that are commonly prescribed to brain tumour patients may have a worsening effect on cognitive function. For example, dexamethasone has been shown to have a negative effect on cognition in a number of different participant populations, including patients with Alzheimer's disease (Weiner et al., 1997); normal elderly participants (Kalmijn et al., 1998) and healthy participants (Kirschbaum et al., 1996). Furthermore, antiepileptic drugs that are commonly prescribed to brain tumour patients who present with seizures have also been found to result in a deterioration on a number of different tests of cognition (Kwan and Brodie, 2001). However, since those patients who were prescribed anti-epileptic medication and/or dexamethasone pre-operatively, were still taking these medications post-operatively, the potential for these medications to negatively affect cognition cannot explain the post-operative deterioration in the brain tumour group compared with baseline performance. Some patients were given a reducing dose of dexamethasone following surgery, compared to their pre-operative dose and therefore any effect of this reduction in dose might have been to improve cognition. For these reasons, the effect of medication is unlikely to explain the observed post-operative deterioration in inspection time and other test performance in the brain tumour group.

The brain tumour and spinal surgery patient groups both showed significant post-operative deterioration on the Edinburgh Functional Impairment Tests by comparison with the healthy control group. The EFITs (i.e. the Nine Hole Peg Tests, Williams Delayed Recall Test and Timed Ten Metre Walk) have previously been shown to be a potentially useful measure of clinical changes in brain tumour patients, and are more sensitive to change than the commonly used disability rating scales, such as the Barthel Disability Index (Grant et al., 1994). The findings of the present study confirm the sensitivity of the EFIT tests to detect treatment related changes following surgery in patients undergoing biopsy or resection of a brain tumour by comparison with a healthy volunteer control group who were simply tested on two separate

occasions. However, in this study the EFIT tests also proved to be a sensitive measure of the effects of spinal surgery since the spinal surgery group also deteriorated post-operatively on these measures, by comparison with the healthy control group. That the WDRT was the only subtest from the EFIT on which the brain tumour group deteriorated significantly more than the matched spinal surgery controls could perhaps be explained by the fact that WDRT performance has been found to be independent of intellectual ability. Grant et al. (1994) suggest that the test is instead actually more sensitive to the structural effects of a brain tumour than are other commonly used cognitive tests that examine digit span or non-verbal learning. The effect size of group in the WDRT general linear model was large, and the brain tumour patients were significantly more impaired at session 2 than either of the control groups. It would therefore appear that the WDRT is potentially as useful as the inspection time task as an index of the effects of brain tumour surgery.

Both surgical groups deteriorated significantly more on the Timed Ten Metre Walk EFIT subtest following surgery than the healthy control group. However, the spinal control group deteriorated to a greater degree than the brain tumour patients following surgery. This finding was expected, given that many spinal surgery patients complained of both pain and an increased level of caution when walking following surgery and it was expected that this would impede the speed with which the Timed Ten Metre Walk was completed. The other tests administered at follow-up that relied heavily on motor function, including digit symbol coding and the Nine Hole Peg Tests, failed to distinguish the two surgical groups since both the brain tumour and spinal surgery groups deteriorated to a similar degree post-operatively. As expected, the healthy control group did not deteriorate on any of these measures when tested on the second occasion. This suggests that surgical intervention itself had a negative effect on motor function generally, as measured by the aforementioned tests. In fact, the extent of the deterioration on these tests that rely on intact motor function may be underestimated in the brain tumour group since a number of the patients in the brain tumour group who did not complete the initial post-operative testing session failed to do so as a result of post-operative complications that included hemiparesis. Had these patients completed session 2

testing following surgery, it is likely that the extent of the observed post-operative deterioration on the tests with a reliance on intact motor function would have been greater than that observed in the group of patients who did complete post-operative follow-up.

Since a number of participants in the brain tumour and spinal surgery groups did not complete session 2 follow-up, detailed comparisons of the session 2 'returners' and 'non-returners' in each surgical group was carried out in order to determine whether those participant who did not complete the initial post-operative follow-up were those who exhibited the most severe impairment at the time of baseline. If this were found to be the case, it could reasonably be assumed that the extent of post-operative deterioration may be underestimated. There were found to be very few significant differences between the demographic characteristics and the baseline test performance of the returners and non-returners in both surgical groups and this therefore suggests that attrition is unlikely to have had any major effect on the results presented in this chapter. The results and discussion of the returners vs. non-returners analyses are discussed in detail in appendix chapter U.

Previous studies examining the utility of the Edinburgh Functional Impairment Tests in brain tumour patients have found there to be a strong relationship between depression, as measured by the Hospital Anxiety and Depression Scale and performance on the speeded EFITs, although the direction of causality of this relationship is not fully understood (Grant et al., 1994). Thus, given that the brain tumour and spinal control groups both reported significantly higher levels of anxiety and depression as assessed by the Hospital Anxiety and Depression Scale at baseline than did the healthy control group, it may be the case that the levels of depression were increased in the two surgical groups post-operatively, following an inpatient hospital stay. If this was the case, increased depression could, in part, explain the deterioration on the speeded tests of functional impairment seen in the brain tumour and spinal surgery groups, but not in the healthy control group. However, no formal measure of mood was made post-operatively. In a previous study of anxiety and depression in patients with supratentorial intracranial tumours in the pre- and post-

operative period, Pringle et al. (1999) found that, compared with assessments made in the pre-operative period, there was a reduction in levels of both anxiety and depression post-operatively. Levels of distress were measured by the Hospital Anxiety and Depression Scale and the findings were similar for both male and female patients, regardless of tumour location and type of surgery (biopsy or tumour resection). The same study also found no significant difference between pre and post-operative Hospital Anxiety and Depression Scale scores in a cohort of patients who underwent elective lumbar spinal surgery. Pringle et al. (1999) conducted their study in the same department and recruited patients from the same hospital wards as in the present study and patients completed the questionnaire prior to discharge following surgery. This reflects the methodology employed in the present study and therefore it is reasonable to assume that a post-operative increase in levels of distress as measured by the Hospital Anxiety and Depression Scale was unlikely and is therefore unable to account for the observed cognitive decline in neither of the surgical patient groups.

It is perhaps more likely that a surgery-related motor impairment could explain the deterioration on these tasks in the brain tumour and spinal control groups. Since inspection time, unlike the other measures, does not rely heavily on motor function and/or a speeded response, and instead provides a measure of visual processing speed, this may explain the lack of relative deterioration on inspection time in the spinal surgery patients in comparison with the brain tumour group. The fact that the brain tumour group performed significantly less well on inspection time following surgery by comparison with both control groups suggests that the task is a sensitive measure of the specific effects of brain tumour surgery on visual information processing. Therefore inspection time is a potentially useful tool with which to detect slowed information processing after surgery that appears to be specific to patients with brain tumours. Thus the findings of the present study suggest that brain tumour surgery has a negative impact on general processing speed in a cohort of patients with a variety of different tumours located in different areas of the brain, when compared with a control group of spinal surgery patients and a healthy control group.

However, the inspection time measure is one that requires a relatively prolonged period of concentration for successful completion, by comparison with the other, shorter tests administered post-operatively (e.g. Digit Symbol Coding and the EFITs). A number of patients in the brain tumour cohort were characterised by reduced concentration and disinhibition at presentation and impaired concentration may also be a result of emotional distress (Taphoorn and Klein, 2004). Thus, these symptoms may not only have influenced performance at baseline but may have worsened following surgery and the effect may be particularly evident on inspection time and the WDRT both of which have comparatively high attentional demands. Therefore, the brain tumour group's deterioration in inspection time performance post-operatively may be the result of reduced concentration as opposed to reflecting an intrinsic impairment in visual processing speed that was significantly worsened by surgical intervention. If this is the case, it may be that post-operative performance of the brain tumour group on any long non-motor cognitive test may have been significantly worse than that of the spinal and healthy control groups.

It is generally accepted that cognitive function requires functional interactions between a number of distinct neural networks located throughout the brain (Bressler, 2002). Studies have found that the presence of a brain tumour resulted in a significant loss of functional connectivity in the brain as measured by MEG analysis, compared with healthy controls and that the loss of connectivity was not only associated with those brain regions located near the tumour, but also in brain regions further away from the tumour itself. These changes in functional connectivity may, at least in part, explain cognitive dysfunction in brain tumour patients (Bartolomei et al., 2006a, Bartolomei et al., 2006b). Studies of the functional anatomy of inspection time using functional connectivity have revealed two networks that are the likely neural correlates of inspection time (see Chapter 2.5). One of these networks includes the fronto-opercular area, intrasylvian area, medial frontal gyrus and the anterior cingulate cortex and the second is a posterior network that includes sensory and associative regions (Deary et al., 2004). The post-surgical deterioration in inspection time scores seen in the brain tumour cohort, but not in either the spinal or healthy control group, may perhaps be the result of surgical intervention further

disrupting functional networks located throughout the brain, particularly those networks that are involved in successful inspection time performance. Future work in this area could involve pre and post-operative functional MRI scanning during inspection time performance in order to determine the functional networks involved in inspection time performance in brain invaded by tumour and to examine the potential role of surgery in disrupting functional connectivity.

Given the heterogeneous nature of the brain tumour group who had a variety of different tumour types, locations and surgical interventions (biopsy or resection), it is of clinical interest to examine the specific effects of these variables on pre and post-operative inspection time and other tests. The results of these analyses are reported in Chapters 7, 8 and 9. Further research should include additional follow-up assessments in order to assess whether or not the initial post-operative deterioration is transient, and whether this differs between different histological groups (see Chapter 9.5 for preliminary analyses).

The cognitive effects of radiotherapy are of particular interest to clinicians and patients alike and have been frequently discussed in the literature (see Chapter 1.5.2). Since inspection time has been prove as a feasible and sensitive measure of brain function in neuro-oncological patients, further research could also involve inspection time testing prior to, during and after radiotherapy as a measure of the effects of this treatment on brain function.

7 The Comparative Effects of Biopsy and Resection

7.1 Overview of analysis procedure

To analyse the effect of biopsy or resection on inspection time, digit symbol-coding and EFIT test performance, general linear modelling (analysis of covariance) was used. The specific test score at session 2 (post-operatively) was specified as the dependent variable. Surgery type (biopsy or resection) and sex were fixed effects (between-groups factors) in the model. Age at the time of testing and NART score were included as covariates for the reasons described in Chapter 5.2.1. The corresponding baseline test score for each participant was also included as an additional covariate in the models. This effectively allows us to compare the change between baseline and session 2 performance on each test between the biopsy and resection groups. There was no significant main effect of sex in any of the models for each of the test scores detailed below. Therefore, sex was omitted as a fixed effect in the analyses reported in this chapter. The models were run again with surgery type (biopsy or resection) specified as the only fixed effect.

In each of the sections that follow, mean baseline and session 2 scores for each surgery group (biopsy and resection) are presented with standard deviations (SD) in brackets. The results of the general linear modelling analysis are then reported. Estimated marginal means (adjusted for age and NART score) are described for each of the tests administered to highlight any differences between the groups, adjusted for other variables in the models.

Of the total cohort of brain tumour patients (n = 118), 30 patients (25.4%) had a biopsy of their tumour and 86 patients (72.9%) had a resection. The remaining 2 patients (1.7%) did not subsequently undergo any surgical intervention and their data are therefore not included in the following analyses.

7.2 Results

7.2.1 Inspection Time Scores: All Inspection Time Data

To examine the effect of surgery (biopsy vs. resection) on inspection time scores, only data from participants who completed inspection time testing at both baseline and post-operative follow-up (session 2) were included in the analysis. When all inspection data (including 'invalid) scores were included, there was data from 15 biopsy patients and 45 resection patients for analysis.

The mean baseline inspection time score for the biopsy group was 113.0 (SD 21.6) and at post-operative follow-up, was 116.3 (SD 22.5). The mean baseline inspection time score for the resection group was 119.2 (SD 20.3) and at post-operative follow-up, was 113.2 (SD 22.7). The mean (standard error) scores at baseline and session 2 for each of the surgery groups are shown in Figure 7.1.

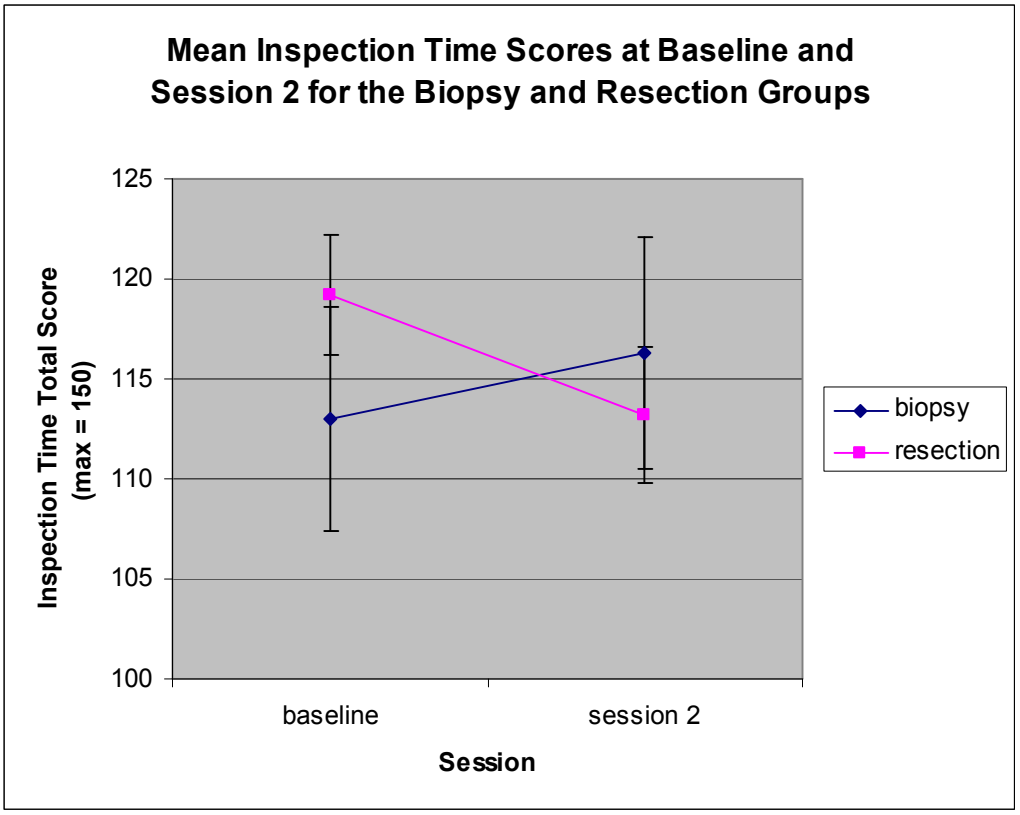


Figure 7.1. Baseline and session 2 inspection time scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with inspection time total score at session 2 entered as the dependent variable, revealed a significant main effect of the inspection time score at baseline, $F(1,55) = 32.484$, $p < 0.001$, partial $\eta^2 = 0.371$. That is, in the whole sample (i.e. the biopsy and resection groups combined), baseline inspection time score was positively correlated with session 2 inspection time score, $r(n = 60) = 0.709$, $p < 0.001$. In the biopsy group, baseline score and session 2 score on inspection time were also significantly correlated, $r(n = 15) = 0.846$, $p < 0.001$. A similar correlation was obtained for the resection group, $r(n = 45) = 0.684$, $p < 0.001$. This therefore shows that those participants with better baseline inspection time scores generally performed significantly better at session 2 follow-up.

The effect of the covariate age on session 2 inspection time scores approached the conventional level of statistical significance, $F(1,55) = 3.772$, $p = 0.057$, partial $\eta^2 = 0.064$. Older participants tended to have lower inspection time scores than younger participants. There was no significant main effect of the covariate NART score in the model, $F(1,55) = 0.621$, $p = 0.434$, partial $\eta^2 = 0.011$. The effect of surgery type did not reach significance in the model that included the effects of the covariates age, NART and baseline inspection time score, $F(1,55) = 3.421$, $p = 0.070$, partial $\eta^2 = 0.059$. That is, in the presence of the effects of age, NART score and baseline inspection time score, there was no significant difference between the performance of the biopsy and resection groups on inspection time at session 2 follow-up.

The estimated marginal mean scores, adjusted for age and NART score, for the biopsy group, were 115.3 (SE 4.6) at baseline, and 120.5 (SE 4.1) at post-operative follow-up. Corresponding mean scores for the resection group were 117.9 (SE 2.6) at baseline, and 111.7 (SE 2.3) at session 2, post-operatively. Thus, despite the fact that the mean scores, shown in Figure 7.1, show that the biopsy group tended to improve after surgery, whereas the resection group deteriorated, this difference between the two surgical groups did not reach statistical significance.

7.2.2 Inspection Time Scores: Valid Inspection Time Data

There were 12 brain tumour patients who had a biopsy with 'valid' inspection time scores at baseline (i.e. at least 17/20 on the longest two durations, see chapter 2.2), who also had follow-up inspection time (valid or invalid) scores at session 2. Forty patients who had tumour resection also met these criteria.

The mean baseline inspection time score for the biopsy group was 120.7 (SD 9.5), and at post-operative follow-up was 120.4 (SD 11.8). For the resection group, the mean baseline score was 124.7 (SD 13.4) and at post-operative follow-up, was 116.6 (SD 21.6). The mean (standard error) scores at baseline and session 2 for the biopsy and resection groups are shown in Figure 7.2.

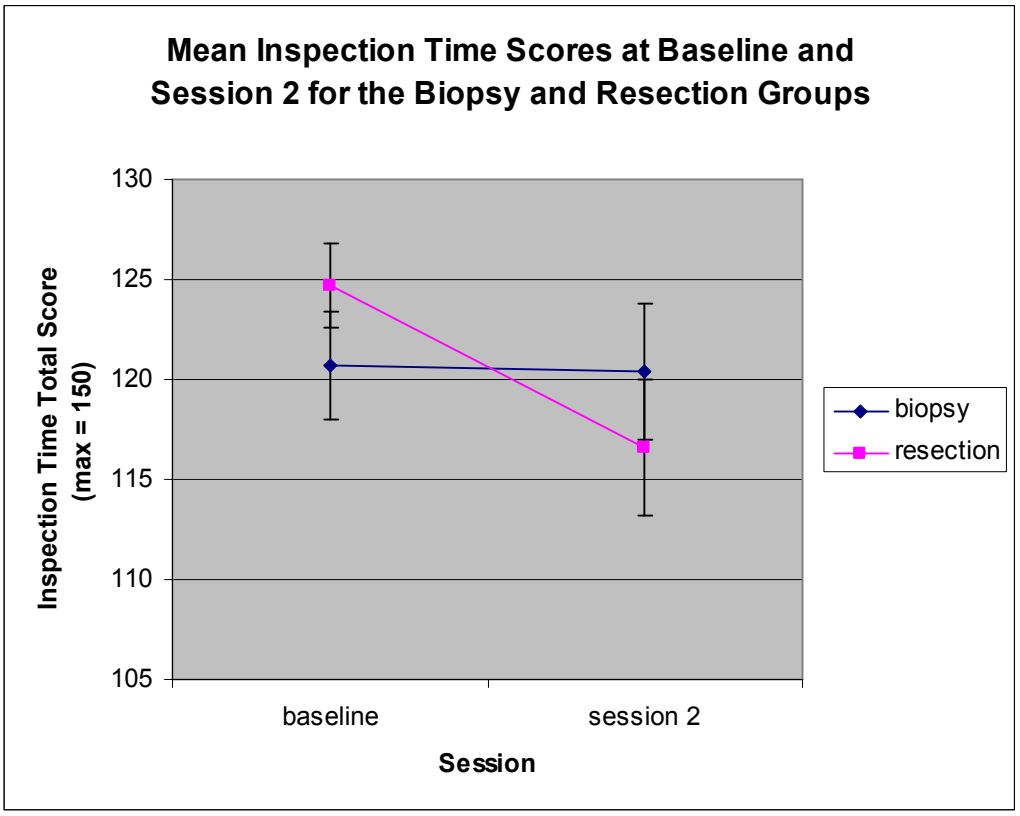


Figure 7.2. Baseline and session 2 inspection time scores for the biopsy and resection groups with only valid baseline inspection scores included. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with inspection time total score at session 2 entered as the dependent variable, revealed a significant main effect of the inspection time score at

baseline, $F(1,47) = 15.294$, $p < 0.001$, partial $\eta^2 = 0.246$. In the whole sample (i.e. the biopsy and resection groups combined), baseline inspection time score was positively correlated with session 2 inspection time score, $r(n = 52) = 0.611$, $p < 0.001$. In the biopsy group, baseline score and session 2 score on inspection time were also significantly correlated, $r(n = 12) = 0.740$, $p = 0.006$. A similar correlation was obtained for the resection group, $r(n = 40) = 0.620$, $p < 0.001$. Those participants with better baseline inspection time scores also generally performed significantly better at session 2 follow-up. There was no significant main effect of the covariate age on session 2 inspection time scores, $F(1,47) = 3.122$, $p = 0.084$, partial $\eta^2 = 0.062$. There was no significant main effect of the covariate NART score in the model, $F(1,47) = 1.349$, $p = 0.251$, partial $\eta^2 = 0.028$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline inspection time score, $F(1,47) = 2.846$, $p = 0.098$, partial $\eta^2 = 0.057$. In the presence of the effects of age, NART score and baseline inspection time score, there was no significant difference between the performance of the biopsy and resection groups on inspection time at session 2 follow-up.

The estimated marginal mean score, adjusted for age and NART score, for the biopsy group at baseline was 122.5 (SE 3.4) and at session 2, post-operatively, was 124.4 (SE 4.6). The corresponding baseline mean for the resection group on the inspection time measure was 124.3 (SE 1.8), and at post-operative follow-up was 115.4 (SE 2.5). Again, despite the fact that the mean scores, shown in Figure 7.2, suggest that when only valid baseline inspection time scores are included in the analysis, the biopsy group maintained their baseline performance when tested post-operatively, whereas the resection group deteriorated following surgery, this difference between the groups was not statistically significant.

7.2.3 Digit Symbol Coding

There were 15 patients who had a biopsy and completed this test at both baseline and post-operative follow-up (session 2). A further 45 patients who had a resection also had pre- and post-operative digit symbol-coding scores.

The mean baseline digit symbol coding score on this test for the biopsy group was 54.3 (SD 18.6), and at session 2 was 53.3 (SD 17.5). The resection group scored a mean of 64.8 (SD 21.7) at baseline and a mean of 59.9 (SD 22.5) at post-operative follow-up. The mean (standard error) scores at baseline and session 2 for each surgery type group are shown in Figure 7.3.

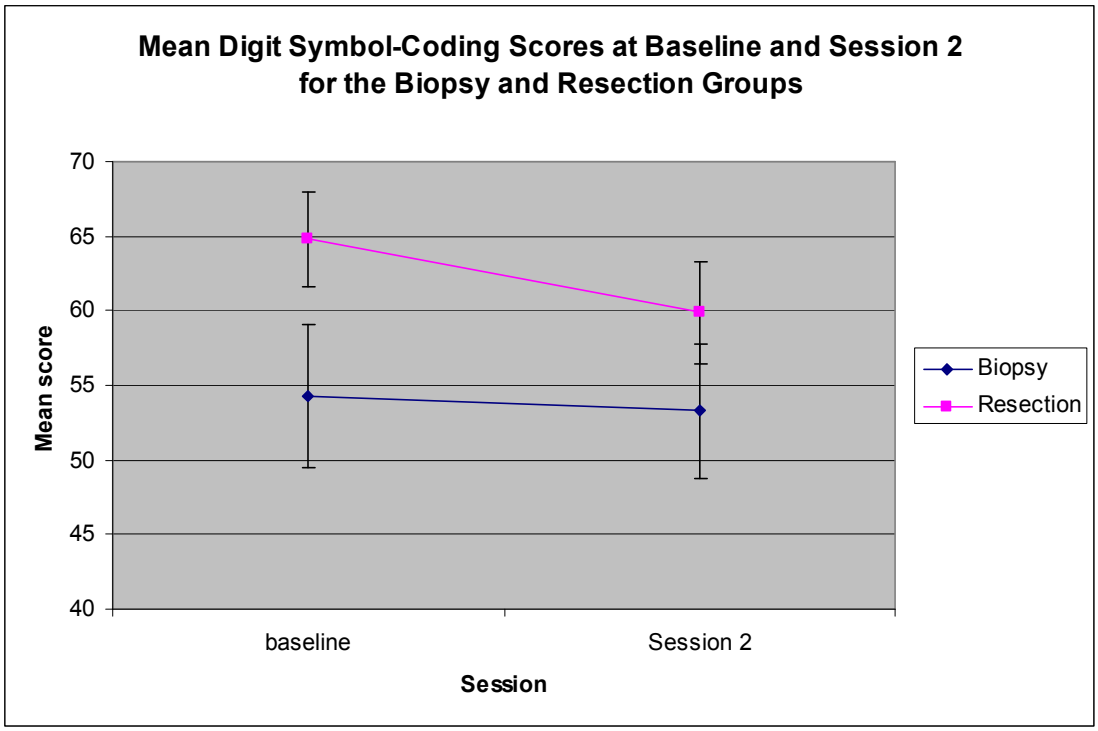


Figure 7.3. Baseline and session 2 digit symbol coding scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

General linear modelling, with digit symbol coding score at session 2 entered as the dependent variable, revealed a significant main effect of the digit symbol coding score at baseline, $F(1,55) = 141.788$, $p < 0.001$, partial $\eta^2 = 0.721$. In the whole sample (i.e. the biopsy and resection groups combined), baseline digit symbol coding score was positively correlated with the corresponding session 2 score, $r(n = 60) = 0.892$, $p < 0.001$. In the biopsy group, baseline score and session 2 score on inspection time were also significantly correlated, $r(n = 15) = 0.953$, $p < 0.001$. A similar correlation was obtained for the resection group, $r(n = 45) = 0.880$, $p < 0.001$. This shows that those participants with better baseline scores on the digit symbol coding test also tended to perform significantly better on the same test at post-

operative session 2 follow-up. The effect of the covariate age on session 2 digit symbol coding scores was not significant $F(1,55) = 0.299$, $p = 0.587$, partial $\eta^2 = 0.005$. There was no significant main effect of the covariate NART score in the model, $F(1,55) = 0.899$, $p = 0.347$, partial $\eta^2 = 0.016$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline inspection time score, $F(1,55) = 0.851$, $p = 0.360$, partial $\eta^2 = 0.015$. In the presence of the effects of age, NART score and baseline digit symbol coding score, there was no significant difference between the performance of the biopsy and resection groups on digit symbol coding at session 2 follow-up

The estimated marginal mean digit symbol coding scores, adjusted for age and NART score, were 56.5 (SE 4.8) at baseline for the biopsy group and 55.3 (SE 4.6) at post-operative follow-up. The resection group had a baseline estimated marginal mean score of 64.0 (SE 2.7) and a mean of 58.8 (SE 2.8) at session 2 follow-up. There was no significant difference in digit symbol coding performance following biopsy or resection of the tumour.

7.2.4 Williams Delayed Recall Test (EFIT)

Fifteen patients who had a biopsy and 47 patients who had a resection completed this EFIT subtest at both baseline and session 2 follow-up.

The biopsy group scored a mean of 10.1 (SD 5.3) at baseline on this measure, and a mean of 12.3 (SD 5.8) at session 2 post-operatively. The resection group had a mean score of 8.3 (SD 5.2) at baseline and a mean of 11.1 (SD 6.2) at session 2 follow-up. The mean (standard error) scores at both baseline and session 2 for the biopsy and resection groups are shown in Figure 7.4.

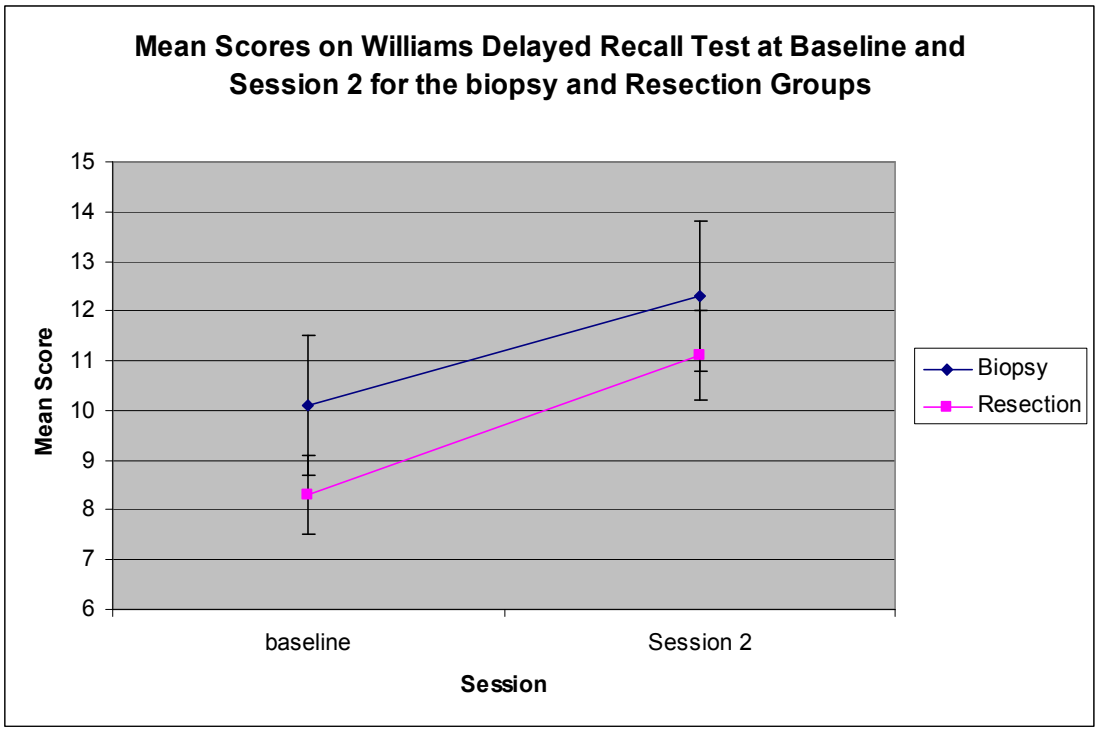


Figure 7.4. Baseline and session 2 Williams Delayed Recall Test scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

Williams Delayed Recall Test (WDRT) score at session 2 was entered as the dependent variable in the general linear model. There was a significant main effect of the test score at baseline, $F(1,57) = 11.288$, $p = 0.001$, partial $\eta^2 = 0.165$. In the whole sample (i.e. the biopsy and resection groups combined), baseline score was positively correlated with the corresponding session 2 score, $r(n = 62) = 0.495$, $p < 0.001$. However, in the biopsy group alone, baseline and session 2 WDRT scores were not significantly correlated, $r(n = 15) = 0.273$, $p = 0.325$. The two scores were significantly correlated in the resection group alone, $r(n = 47) = 0.552$, $p < 0.001$. This therefore shows that, overall, those participants with better baseline scores on the WDRT also performed significantly better on the same test at post-operative session 2 follow-up. The effect of the covariate age on session 2 WDRT scores was not significant $F(1,57) = 1.103$, $p = 0.298$, partial $\eta^2 = 0.019$. There was no significant main effect of the covariate NART score in the model, $F(1,57) = 0.538$, $p = 0.466$, partial $\eta^2 = 0.009$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline inspection time score, $F(1,57) = 0.006$, $p = 0.937$, partial $\eta^2 < 0.001$. In the presence of the

effects of age, NART score and baseline WDRT score, there was no significant difference between the performance of the biopsy and resection groups on the WDRT at session 2 follow-up.

The estimated marginal mean score, adjusted for age and NART score for the biopsy group at baseline was 9.6 (SE 1.3), compared with 11.5 (SE 1.4) at post-operative testing. The corresponding means for the resection group were 8.5 (SE 0.7) and 11.4 (SE 0.8), respectively. Figure 7.4 demonstrates that both groups deteriorated following surgery, but that the extent of this deterioration did not differ between the biopsy and resection groups.

7.2.5 Nine Hole Peg Test (Right Hand, EFIT)

Fifteen patients who had a biopsy of their tumour completed this measure at both baseline and session 2 and 47 patients who underwent a resection also completed the task both pre- and post-operatively.

The biopsy group scored a mean of 14.4 (SD 3.4) at baseline, and 15.3 (SD 3.7) at session 2, post-operatively. The resection group scored a mean of 13.6 (SD 2.8) at baseline and 14.3 (SD 2.5) at post-operative testing. The mean (standard error) scores at baseline and session 2 for the biopsy and resection groups are shown in Figure 7.5.

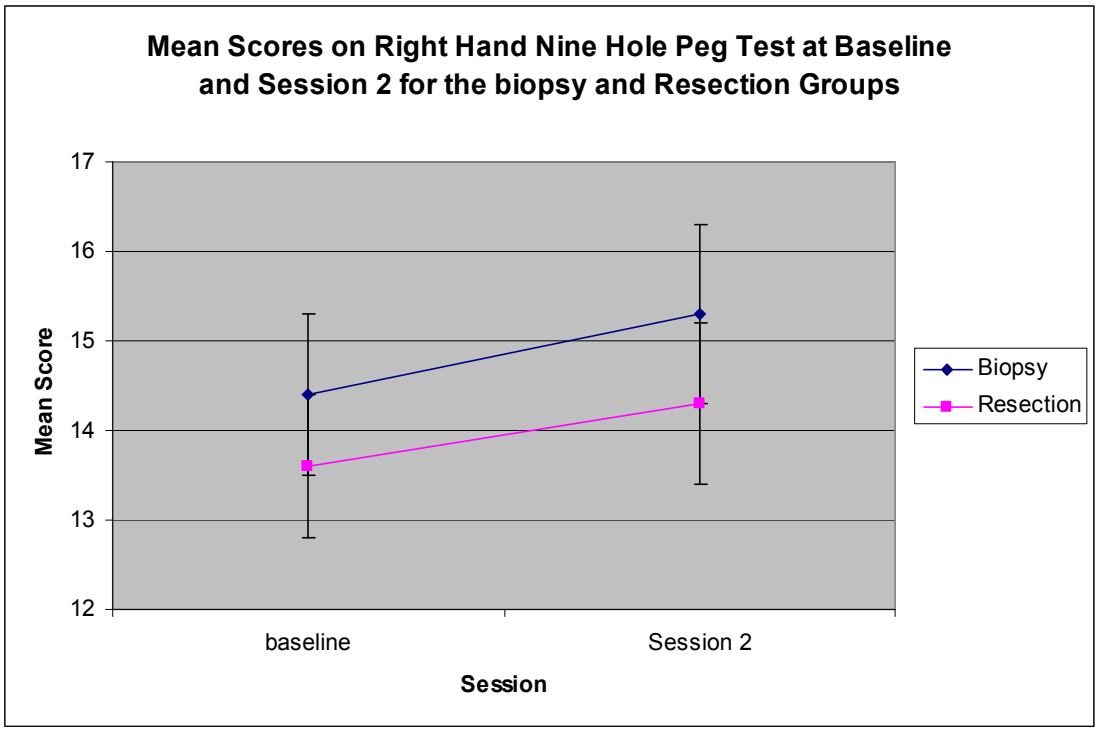


Figure 7.5. Baseline and session 2 right hand nine hole peg test scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

Right hand nine hole peg test score at session 2 was entered as the dependent variable in the general linear model. There was a significant main effect of the test score at baseline, $F(1,57) = 45.174$, $p < 0.001$, partial $\eta^2 = 0.442$. In the whole sample (i.e. the biopsy and resection groups combined), baseline score was positively correlated with the corresponding session 2 score, $r(n = 62) = 0.741$, $p < 0.001$. In the biopsy group, the baseline score and session 2 score were also significantly correlated, $r(n = 15) = 0.833$, $p < 0.001$. The two scores were also significantly correlated in the resection group, $r(n = 47) = 0.692$, $p < 0.001$. This therefore shows that those participants with better baseline scores on the digit symbol coding test also performed significantly better on the same test at post-operative session 2 follow-up. The effect of the covariate age on session 2 right hand nine hole peg test scores was not significant $F(1,57) = 0.728$, $p = 0.397$, partial $\eta^2 = 0.013$. There was no significant main effect of the covariate NART score in the model, $F(1,57) = 0.151$, $p = 0.699$, partial $\eta^2 = 0.003$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline score, $F(1,57) = 1.225$, $p = 0.273$, partial $\eta^2 < 0.021$. In the presence of the effects of age,

NART score and baseline inspection time score, there was no significant difference between the performance of the biopsy and resection groups on the right hand nine hole peg test at session 2 follow-up

The estimated marginal mean scores, adjusted for age and NART score, were 14.1 (SE 0.6) for the biopsy group at baseline and were 15.0 (SE 0.5) at session 2, post-operatively. The corresponding means for the resection group were 13.7 (SE 0.3) at baseline and 14.4 (SE 0.3) at session 2. Figure 7.5 shows that both the biopsy and resection group were slower at post-operative testing compared with baseline. There was no significant difference in the extent of decline between the groups.

7.2.6 Nine Hole Peg Test (Left Hand, EFIT)

Fifteen biopsy patients and 47 resection patients completed this test at both baseline and post-operative follow-up (session 2).

The biopsy group had a mean score of 15.6 (SD 3.4) at baseline testing, and 16.0 (SD 3.9) at session 2. The resection group had a mean score of 15.2 (SD 3.7) at baseline and 15.8 (SD 3.4) post-operatively at session 2. The mean (standard error) scores at baseline and session 2 for the biopsy and resection groups are shown in Figure 7.6.

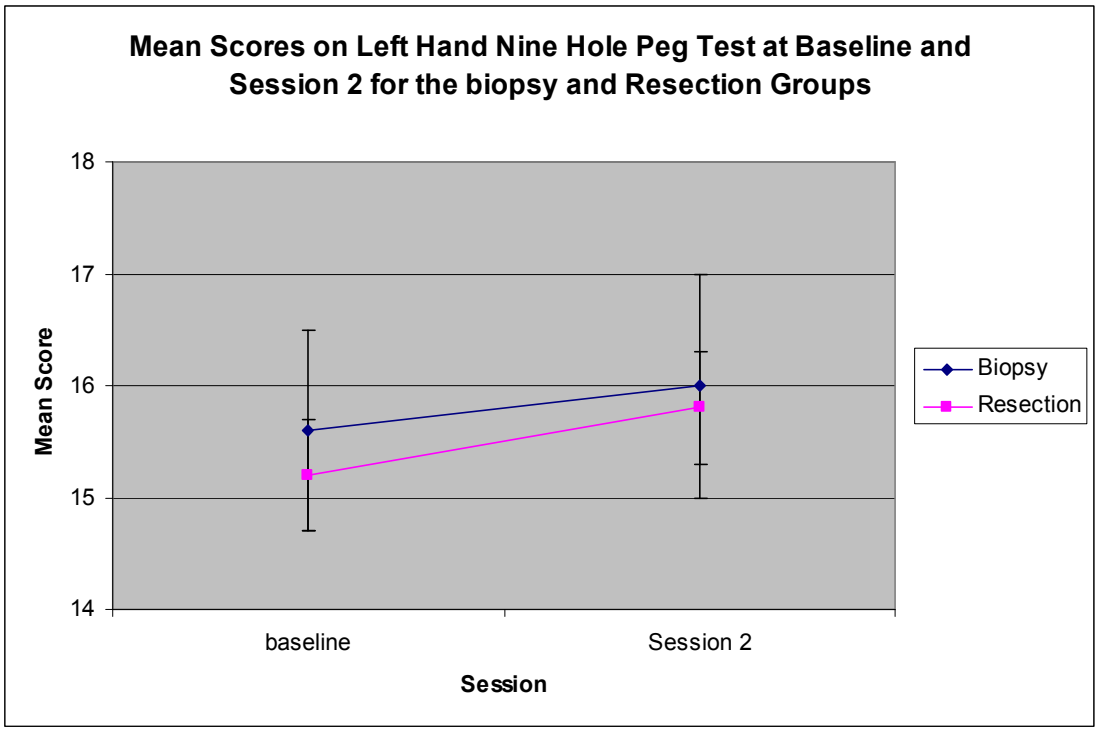


Figure 7.6. Baseline and session 2 left hand nine hole peg test scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

Left hand nine hole peg test score at session 2 was entered as the dependent variable in the general linear model. There was a significant main effect of the test score at baseline, $F(1,56) = 125.546$, $p < 0.001$, partial $\eta^2 = 0.692$. In the whole sample (i.e. the biopsy and resection groups combined), baseline score was positively correlated with the corresponding session 2 score, $r(n = 61) = 0.859$, $p < 0.001$. In the biopsy group, baseline score and session 2 score on the left hand nine hole peg test were also significantly correlated, $r(n = 15) = 0.797$, $p < 0.001$. The two scores were also significantly correlated in the resection group, $r(n = 46) = 0.886$, $p < 0.001$. This therefore shows that those participants with better baseline scores on the left hand nine hole peg test also performed significantly better on the same test at post-operative session 2 follow-up. The effect of the covariate age on session 2 scores was significant $F(1,56) = 6.945$, $p = 0.011$, partial $\eta^2 = 0.110$. Older participants took significantly longer to complete the test than younger participants. There was no significant main effect of the covariate NART score in the model, $F(1,56) = 0.553$, $p = 0.460$, partial $\eta^2 = 0.010$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline inspection

time score, $F(1,57) = 0.355$, $p = 0.554$, partial $\eta^2 < 0.006$. In the presence of the effects of age, NART score and baseline score, there was no significant difference between the performance of the biopsy and resection groups on the right hand nine hole peg test at session 2 follow-up.

The estimated marginal mean score, adjusted for age and NART score, for the biopsy group was 15.3 (SE 0.9) at baseline and was 15.6 (SE 0.5) at session 2 testing. The estimated marginal mean score for the resection group was 15.3 (SE 0.5) at baseline and 15.9 (SE 0.3) at session 2 follow-up. Figure 7.6 shows that neither the biopsy nor the resection group deteriorated significantly between baseline and session 2 testing on this measure.

7.2.7 Timed Ten Metre Walk (EFIT)

Fourteen patients in the biopsy group completed this measure at both baseline and session 2, post-operatively and 42 patients in the resection group also completed the measure on both occasions.

The mean score for the biopsy group at baseline was 6.8 (SD 1.2) and post-operatively was 7.2 (SD 1.6). The resection group had a mean score of 6.7 (SD 1.5) at baseline, and a mean of 7.1 (SD 1.7) at session 2 follow-up. The mean (standard error) scores at baseline and session 2 for the biopsy and resection groups are shown in Figure 7.7.

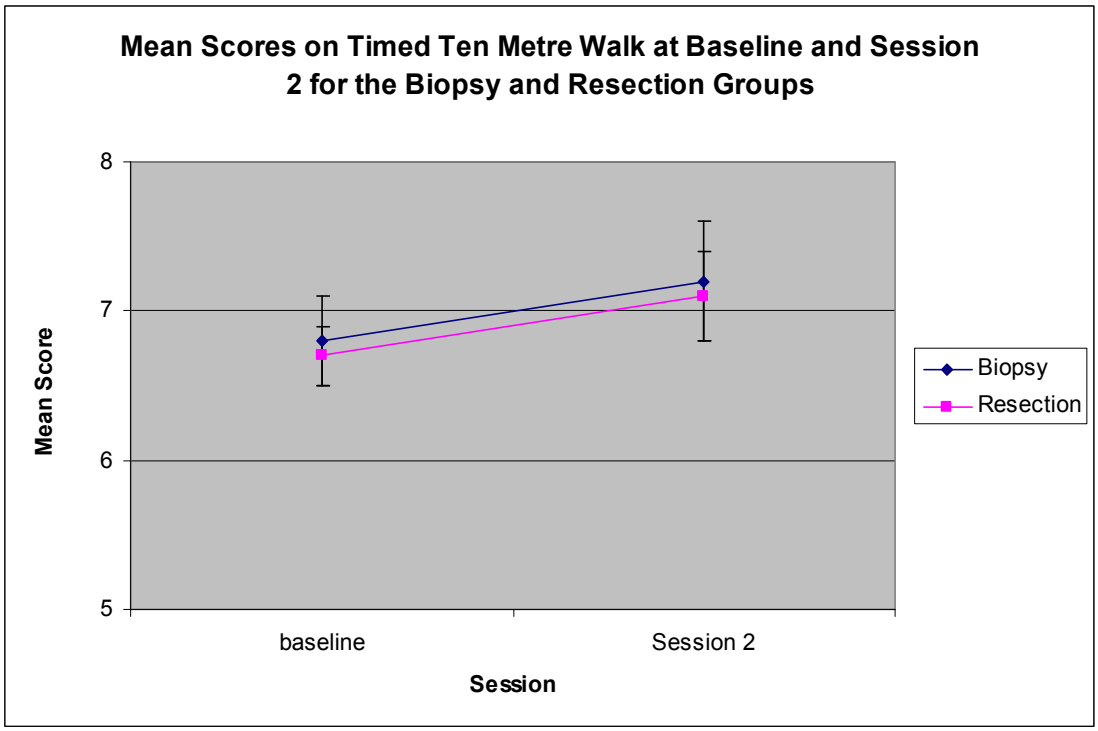


Figure 7.7. Baseline and session 2 timed ten metre walk test scores for the biopsy and resection groups. Points are raw mean scores. Error bars: Standard error means.

Timed ten metre walk test score at session 2 was entered as the dependent variable in the general linear model. There was a significant main effect of the test score at baseline, $F(1,51) = 36.855$, $p < 0.001$, partial $\eta^2 = 0.419$. In the whole sample (i.e. the biopsy and resection groups combined), baseline timed ten metre walk score was positively correlated with the corresponding session 2 score, $r(n = 56) = 0.731$, $p < 0.001$. In the biopsy group, baseline score and session 2 score on this test were also significantly correlated, $r(n = 14) = 0.781$, $p = 0.001$. The correlation between the two scores in the resection group was also significant, $r(n = 42) = 0.721$, $p < 0.001$. This therefore shows that those participants who completed the timed ten metre walk fastest at baseline also performed better on the same test at post-operative session 2 follow-up. The effect of the covariate age on session 2 scores was significant $F(1,51) = 4.213$, $p = 0.045$, partial $\eta^2 = 0.076$. Older participants took significantly longer to complete the test than younger participants. There was no significant main effect of the covariate NART score in the model, $F(1,51) = 0.350$, $p = 0.557$, partial $\eta^2 = 0.007$. The effect of surgery type was not significant in the model that included the effects of the covariates age, NART and baseline timed ten metre walk score, $F(1,51)$

= 0.134, p = 0.716, partial η^2 = 0.003. In the presence of the effects of age, NART score and baseline inspection time score, there was no significant difference between the performance of the biopsy and resection groups on the time ten metre walk test at session 2 follow-up.

The estimated marginal mean scores, adjusted for age and NART effects, were 6.5 (SD 0.4) for the biopsy group at baseline, and were 7.0 (SD 0.3) at post-operative follow-up. The estimated marginal means scores for the resection group were 6.9 (SD 0.2) at baseline, and 7.1 (SD 0.2) at follow-up. These means, together with Figure 7.7 shows that neither the biopsy or resection group deteriorated significantly between baseline and post-surgical follow-up.

7.3 Discussion

Comparison of those patients who had a tumour biopsy and those who underwent tumour resection revealed no significant differences in terms of the extent of post-operative deterioration on any of the tests administered at session 2. However, there was a clear trend towards maintenance of pre-operative performance in the biopsy group and post-operative deterioration in the resection group on the inspection time task, although the difference failed to reach the conventional level of statistical significance.

The general cognitive effects of surgical intervention in brain tumour patients are discussed in chapter 6.3. Thus, only the specific effects of tumour biopsy vs. resection are discussed here.

This is one of the first studies to specifically compare the cognitive effects of biopsy and resection. There was a high rate of attrition in the brain tumour group which resulted in only 54% of the brain tumour cohort participating in the initial post-operative follow-up testing session (session 2). Therefore, there were only 16 patients in the biopsy group and 48 patients in the resection group who had session 2 follow-up scores and some of these patients did not complete the entire follow-up test battery. As a result, the power to detect any differences in the post-operative response of the two brain tumour surgery groups is considerably reduced, making a type 2 statistical error likely. This lack of power may explain why, although a clear trend towards there being a deleterious effect of resective surgery on inspection time performance compared with the lesser effects of biopsy, the difference in performance of the two groups did not reach statistical significance. Had a greater number of patients completed the post-operative test battery at session 2, the difference between the biopsy and resection groups may have been larger and statistically significant. That there may have been some bias that resulted in the better functioning patients completing follow-up testing must also be considered. However, the proportion of patients in each surgery group who did complete session 2 post-operative follow-up was similar (53% of the biopsy group and 56% of the

resection group). This may suggest that a similar proportion of patients in the biopsy and tumour groups either declined, were missed or too unwell to take part in the initial post-operative testing session, and thus reduces the potential for bias from the drop-outs in this instance.

Very few studies have addressed the question of whether tumour biopsy and resection have differential effects on the cognitive function of brain tumour patients. This question is of interest given the limited survival times of many patients with high-grade gliomas in particular. Should resection be proven to increase survival times, it is of interest to determine whether this may be at the expense of compromised cognitive function and thus likely impairment to quality of life. In the present study, patients were classified simply into biopsy or resection groups. Any tumour operation where more tissue than necessary for a histopathological diagnosis was removed was deemed to be a 'resection' (Hart et al., 2000). The potential alternative method of classifying patients into more specific groups such as 'partial resection' and gross total resection' would have been problematic as this would have relied solely on the judgement of the neurosurgeon. Given that a number of different neurosurgeons carry out brain tumour surgery in our department, there would have been considerable individual differences in terms of assessment regarding the extent of resection. Moreover, more specific classification with regards to type of surgery would have further reduced the power to detect any differences between the groups.

The potential survival benefits of tumour resection as opposed to biopsy in patients with gliomas have received increasing attention in recent years (Smith et al., 2008, Stummer et al., 2008). However, whilst in high grade gliomas the evidence is derived from a randomised clinical trial (Stummer et al., 2008), answering this question has proven to be much more problematic in low grade gliomas given that those patients who undergo tumour resection tend to be selected for smaller tumours, more accessible tumours and are more likely to be reported than a conservatively managed case series.

Studies have also evaluated the neurosurgical morbidity and mortality associated with biopsy and resection procedures. It is generally accepted that tumour biopsy is a relatively lower risk procedure with a low morbidity and mortality rate (Hall, 1998). Reported mortality rates and morbidity rates tend to be higher for tumour resections (Sawaya et al., 1998). However, Vecht et al. (1990) examined neurological function in malignant glioma patients in relation to the extent of surgery. Large resections were found to not necessarily result in more neurological impairment than more limited surgery. They report that where more extensive surgery was carried out, this actually tended to improve the neurological function of the patient. However the closer the resection to eloquent brain regions (e.g. somatosensory regions, language areas) the more likely a larger resection is to cause a deficit (Peraud et al., 2002).

Therefore, the potentially beneficial effects of tumour resection in terms of increasing survival time and the potential for increased functional deficits following extensive resection compared with biopsy have received the greatest attention in this area. The cognitive effects of biopsy compared to resection are an understudied area and future studies could compare inspection time, and other test performance between the two groups in a larger cohort than the present study, and including more long-term follow-up assessments. This would allow us to determine whether inspection time is a sensitive measure of the specific deleterious effects of tumour resection by comparison with biopsy.

8 Tumour Type, Location and Lateralisation: Baseline/Pre-Operative Function

8.1 Overview of Analysis Procedure

To examine the effect of the following tumour-related variables: histological subtype, location (lobe) and hemispheric lateralisation, on baseline inspection time scores (i.e. prior to any surgical intervention), general linear modelling (ANCOVA) was used. The relevant tumour-related variable (i.e. type, lobe or hemisphere) and sex were entered as fixed factors in the model. Age and National Adult Reading Test (NART) score were entered as covariates for the reasons detailed in Chapter 5.2.1. Baseline inspection time score was the dependent variable in each analysis. The tumour-related variable by sex interaction was never significant and is therefore not reported in the following analyses. Least Significant Difference (LSD) pairwise comparisons were specified to examine any differences between subgroups and any significant comparisons are reported, where applicable.

8.2 Histological Tumour Type

The brain tumour patients were classified into one of the following five main groups, according to the histological classification of their brain tumour:

(i) Low-grade glioma (WHO grades I or II)
(ii) High-grade glioma (WHO grades III or IV)
(iii) Metastasis
(iv) Meningioma
(v) Other (including pituitary adenomas)

The histological diagnoses were obtained from the neuropathology department following surgical intervention. Of the total cohort of brain tumour patients (n = 118), 23 patients were found to have a low-grade glioma. Fifty-one patients had a high-grade glioma and 14 had a metastatic brain tumour that had originated from a primary cancer elsewhere in the body. A further 17 had a meningioma and 11 patients had a brain tumour that did not fall into any of the aforementioned categories ('other'). The remaining 2 patients were those patients for whom surgery was initially planned but did not subsequently take place. Since six patients did not complete the inspection time measure at baseline, the high-grade glioma group fell to a total of 46 patients and the meningioma group to 16 patients for this analysis.

In the following analyses where a significant main effect of tumour type was not found, post-hoc analyses were carried out to compare the low-grade and high-grade tumour type groups only, since it was hypothesised that the performance of these two groups was most likely to be significantly different (see Chapter 1.3).

8.2.1 Inspection Time Scores: All Inspection Time Data

Twenty-three low-grade glioma patients, 46 high-grade glioma patients, 14 patients with a metastasis, 16 with a meningioma and 11 with a tumour classified as 'other' all completed the inspection time task at baseline.

All inspection time data, including 'invalid' scores (see chapter 2.2), were entered into the model in the first instance. When all data was included, the group of low-grade glioma patients scored a mean of 126.1 (SD 11.5) on the task at baseline (i.e. pre-operatively). The high-grade glioma group had a mean score of 110.2 (SD 22.2) and the group of patients with metastatic tumours scored a mean of 113.8 (SD 18.3). The meningioma patients had a mean of 106.4 (SD 22.6) and the group with tumours classified as 'other' scored a mean of 119.2 (SD 11.6) on the inspection time measure at baseline.

The covariate 'age' had a significant main effect in the general linear model, $F(1,98) = 26.197$, $p < 0.001$, partial $\eta^2 = 0.211$. The covariate NART score also had a significant main effect in the model, $F(1,98) = 5.708$, $p = 0.019$, partial $\eta^2 = 0.055$. Older patients and patients with higher (poorer) NART scores were more likely to have lower (poorer) inspection time scores. Sex had no significant main effect, $F(1,98) = 2.527$, $p = 0.115$, partial $\eta^2 = 0.025$. Tumour type also had no significant main effect in the model that included the effects of age and NART, $F(4,98) = 1.280$, $p = 0.283$, partial $\eta^2 = 0.050$. Therefore, the type of tumour had no overall effect on inspection time scores at baseline.

The estimated marginal mean inspection time scores, adjusted for age and NART score and shown in Figure 8.1, for each tumour type were 120.6 (SE 3.8) for the low-grade glioma group; 111.0 (SE 2.5) for the high-grade glioma group; 118.0 (SE 4.7) for the metastasis group; 115.4 (SE 8.9) for the meningioma group and 114.7 (SE 5.4) for the group of patients with 'other' tumours. Pairwise comparisons using Least Significant Difference (LSD) tests showed that the high-grade glioma group performed significantly worse than the low-grade glioma group ($p = 0.040$).

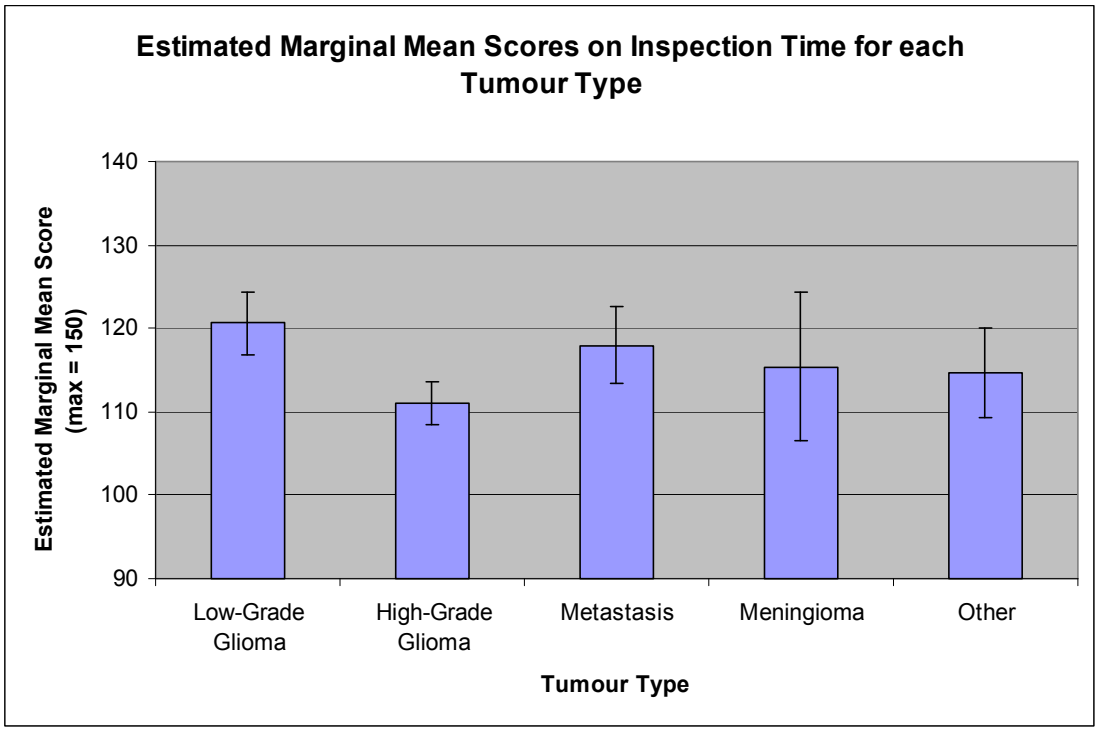


Figure 8.1. Baseline inspection time scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.2 Inspection Time Scores: Valid Inspection Time Data

The previous analysis was repeated including only 'valid' inspection time scores (i.e. data from patients who scored ≥ 17/20 on the longest two durations on the task; see Chapter 2.2). This resulted in data from 23 low-grade glioma patients, 37 high-grade glioma patients, 12 patients with metastatic tumours, 13 patients with a meningioma and 9 patients with tumours classified as 'other', being included in the analysis.

All low-grade glioma patients had valid inspection time date at baseline, thus the mean score for this group remained 126.1 (SD 11.5). The high-grade glioma group had a mean score of 118.5 (SD 14.2) when only valid inspection time data was analysed. The corresponding mean for the metastasis group was 118.7 (SD 14.7) and, for the group of patients with a meningioma, was 114.5 (SD 15.8). The patients with a tumour type classified as 'other' had a mean score of 123.1 (SD 8.0) in this instance.

The covariate 'age' again had a significant main effect in the model, $F(1,82) = 15.627$, $p < 0.001$, partial $\eta^2 = 0.160$. Older patients had lower inspection time scores than younger patients. The covariate NART was not significant on this occasion, $F(1,82) = 1.821$, $p = 0.181$, partial $\eta^2 = 0.022$. Sex had a significant effect in this model, $F(1,82) = 4.157$, $p = 0.045$, partial $\eta^2 = 0.048$. Female patients tended to have lower scores than male patients. Again, tumour type had no significant main effect in the model that included the effects of the covariates age and NART score, $F(4,82) = 0.683$, $p = 0.605$, partial $\eta^2 = 0.032$.

The estimated marginal mean scores for the valid inspection time score, adjusted for age and NART score and shown in Figure 8.2, for each tumour type were 123.7 (SE 2.7) for the low-grade glioma group; 118.4 (SE 2.0) for the high-grade glioma group; 122.3 (SE 3.7) for the metastasis group; 121.6 (SE 6.5) for the meningioma group and 120.6 (SE 4.3) for the 'other' tumour group. Pairwise comparisons using LSD

tests found no significant difference between the performance of the low-grade and high-grade glioma groups.

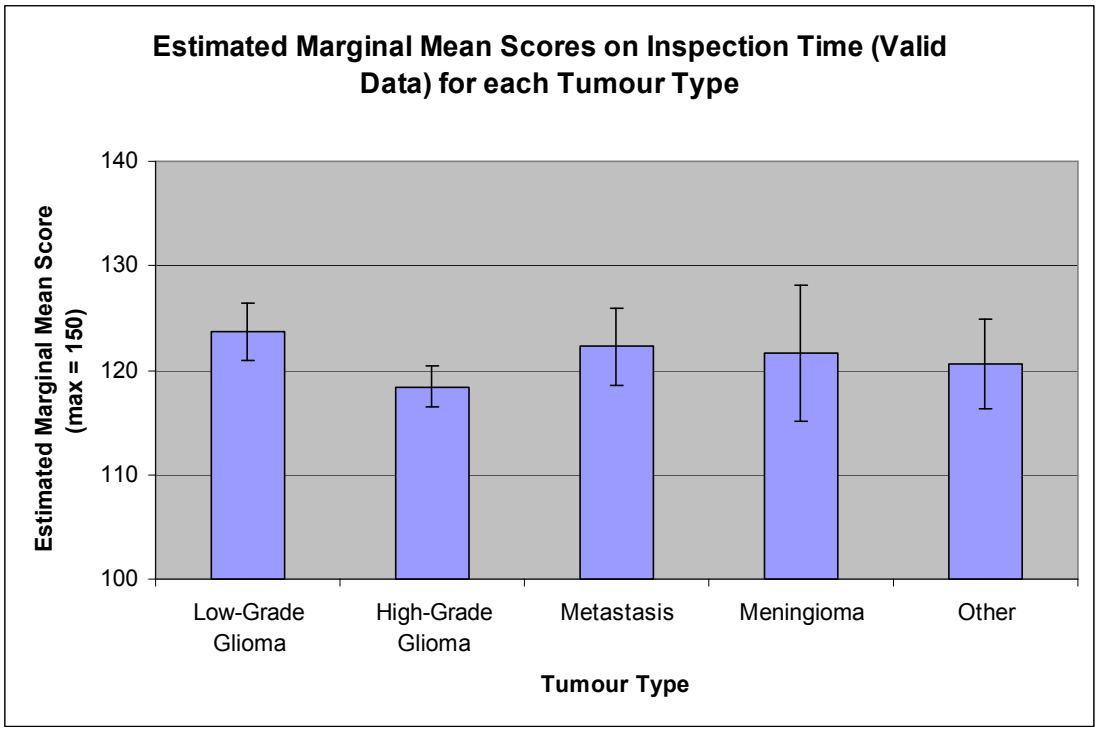


Figure 8.2. Baseline inspection time scores including only valid data for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means

8.2.3 Rey Auditory Verbal Learning Test

There were 20 patients with low-grade gliomas, 35 with high-grade gliomas, 10 with a metastasis, 10 with a meningioma and 10 with 'other' tumour types who completed the Rey Auditory Verbal Learning Test at baseline.

The mean score for the low-grade glioma group on this test was 65.5 (SD 16.1) and for the high-grade glioma group was 56.3 (SD 19.1). The metastasis group had a mean score of 65.3 (SD 17.2) and the meningioma group had a mean score of 66.5 (SD 13.0). The group of patients with 'other' tumour types scored a mean of 65.3 (SD 18.2).

There was a significant main effect of the covariates age, $F(1,73) = 11.899$, $p = 0.001$, partial $\eta^2 = 0.140$; and NART score, $F(1,73) = 16.141$, $p < 0.001$, partial $\eta^2 =$

0.181, in the model. Younger patients and patients with lower (better) NART scores were more likely to recall more words on this measure. The effect of sex was not significant in the model, $F(1,73) = 1.409$, $p = 0.239$, partial $\eta^2 = 0.019$. The effect of tumour type on RAVLT scores did not reach the conventional level of statistical significance in the model that included the effects of age and NART score, $F(4,73) = 2.259$, $p = 0.071$, partial $\eta^2 = 0.110$.

The estimated marginal mean scores on the RAVLT, adjusted for age and NART score and shown in Figure 8.3, for each tumour type were 62.3 (SE 3.8) for the low-grade glioma group; 57.8 (SE 2.5) for the high-grade glioma group; 61.6 (SE 5.1) for the metastasis group; 76.9 (SE 8.1) for the meningioma group and 71.5 (SE 5.2) for the 'other' tumour group. Pairwise comparisons using LSD tests revealed no significant differences between the low-grade and high-grade glioma groups.

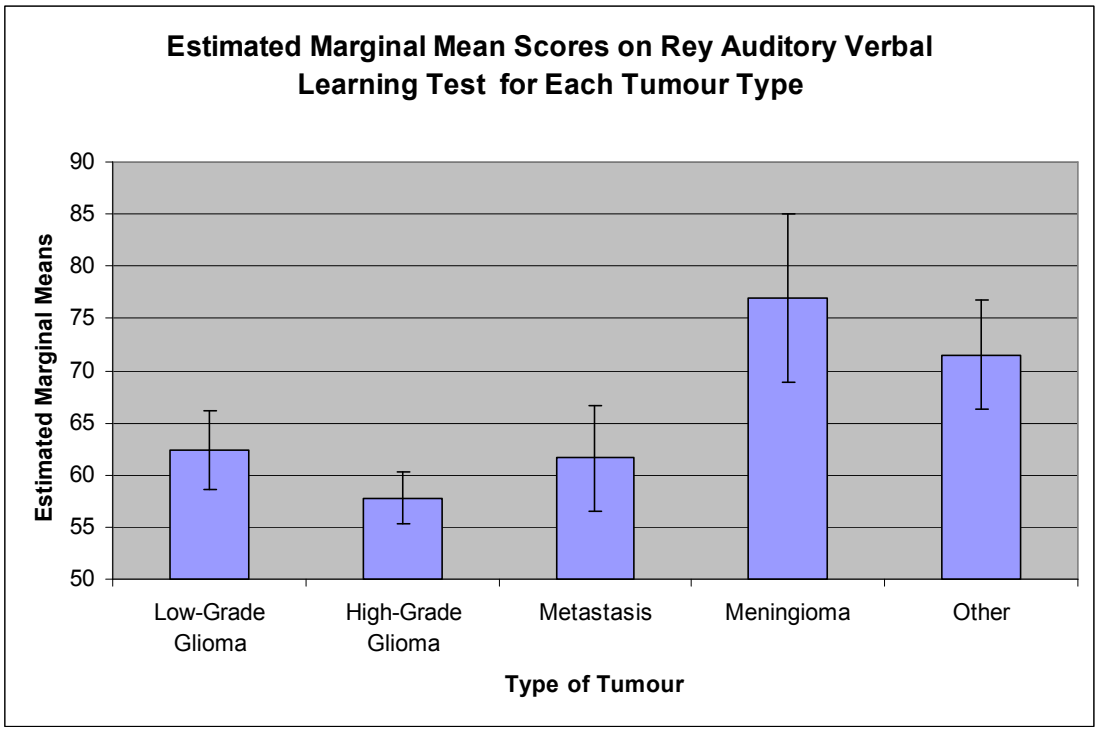


Figure 8.3. Baseline Rey Auditory Verbal Learning Test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means

8.2.4 Trail Making Test Part B

Twenty low-grade glioma patients, 41 high-grade glioma patients, 11 metastasis patients, 12 meningioma and 10 patients with 'other' tumours completed the trail making test part B at baseline.

The mean score for the low-grade glioma group on this test was 83.1 (SD 21.5). The high-grade glioma group scored a mean of 100.2 (SD 32.9), the metastasis group scored a mean of 91.9 (SD 33.6) and the meningioma group scored a mean of 99.7 (SD 34.8). The group of patients with 'other' tumours scored a mean of 85.2 (SD 29.0) on the test at baseline.

There was a significant main effect of both covariates age, $F(1,82) = 13.861$, $p < 0.001$, partial $\eta^2 = 0.145$; and NART score, $F(1,82) = 21.141$, $p < 0.001$, partial $\eta^2 = 0.205$, in the model. Older patients and patients who had higher (poorer) NART scores took significantly longer to complete the task. The effect of sex was not significant, $F(1,82) = 0.001$, $p = 0.970$, partial $\eta^2 < 0.001$. Tumour type had no significant main effect on trail making test part B score in the model that included the effects of age and NART score, $F(4,82) = 0.830$, $p = 0.510$, partial $\eta^2 = 0.039$.

The estimated marginal mean scores on the trail making test part B, adjusted for age and NART score and shown in Figure 8.4, were 90.2 (SE 6.5) for the low-grade glioma group; 99.1 (SE 4.2) for the high-grade glioma group; 93.8 (SE 8.7) for the metastasis group; 105.6 (SE 14.4) for the meningioma group and 85.4 (SE 8.7) for the group of patients with 'other' tumours. Pairwise comparisons revealed no significant difference between the performance of the low-grade and high-grade glioma groups.

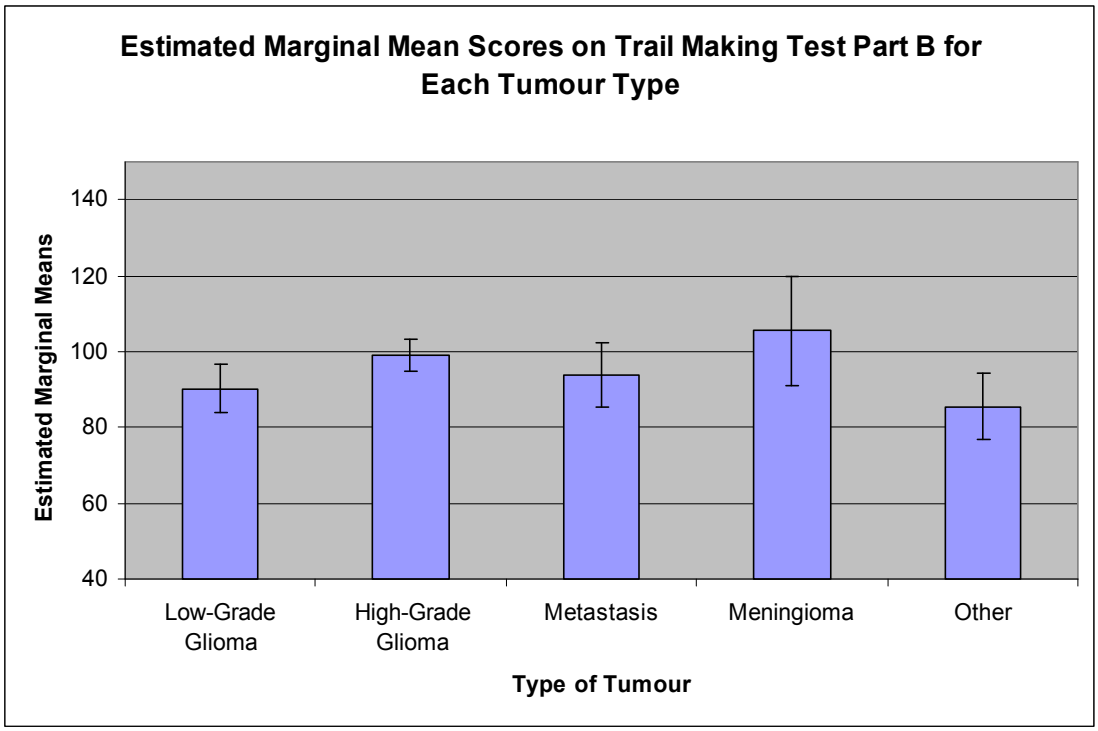


Figure 8.4. Baseline trail making test part B scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.5 Verbal Fluency

Eighteen patients with a low-grade glioma, 39 with a high-grade glioma, 11 metastasis patients, 11 meningioma patients and 10 patients with 'other' tumours all completed the verbal fluency test at baseline.

The low-grade glioma group had a mean score of 34.9 (SD 14.9) on this measure. The high-grade glioma group had a mean of 23.9 (SD 8.9), the metastasis group had a mean of 37.3 (SD015.7), the meningioma group mean score was 26.4 (SD7.0) and the 'other' tumour group scored a mean of 29.7 (SD 9.1) on verbal fluency at baseline.

The effect of the covariate age was not significant in the model, $F(1,77) = 0.767$, $p = 0.384$, partial $\eta^2 = 0.010$. The covariate NART did have a significant main effect in the model, $F(1,77) = 22.581$, $p < 0.001$, partial $\eta^2 = 0.227$, with patients who had lower (better) NART scores more likely to produce more words on the verbal fluency

task. Sex had no significant effect in the model, $F(1,77) = 1.663$, $p = 0.201$, partial $\eta^2 = 0.021$. The effect of tumour type was significant in the model that included the effects of age and NART score, $F(4,77) = 5.700$, $p < 0.001$, partial $\eta^2 = 0.228$.

The estimated marginal mean verbal fluency scores, adjusted for age and NART score and shown in Figure 8.5, for each tumour group at baseline were 35.8 (SE 2.3) for the low-grade glioma group; 24.1 (SE 1.5) for the high-grade glioma group; 33.3 (SE 2.9) for the metastasis group; 30.7 (SE 5.0) for the meningioma group and 30.7 (SE 3.0) for the group of 'other' tumour patients. Since there was a significant main effect of tumour type in the model, post-hoc analyses using LSD pairwise comparisons were carried out comparing the performance of each of the tumour type groups. These showed that the high-grade glioma group performed significantly less well than both the low-grade glioma group ($p < 0.001$) and the metastasis group ($p = 0.006$).

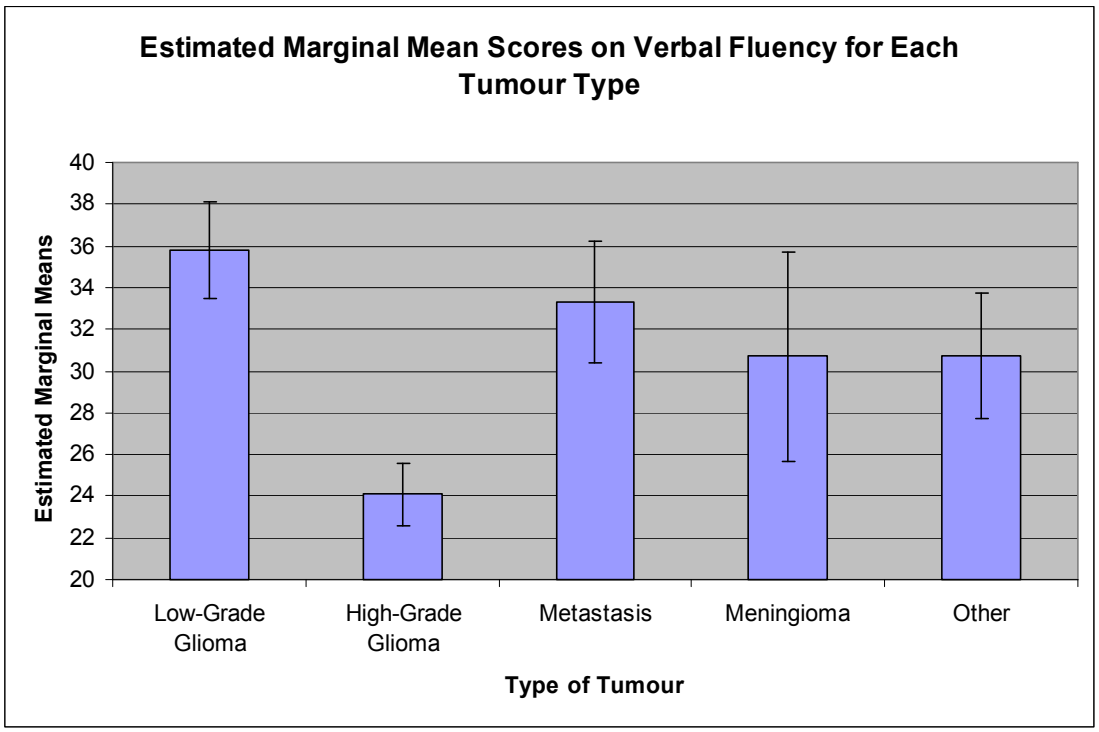


Figure 8.5. Baseline verbal fluency test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.6 Digit Symbol Coding

Twenty-two members of the low-grade glioma group completed digit symbol coding at baseline. Forty-six high-grade glioma patients, 14 metastasis patients, 17 meningioma patients and 10 'other' tumour patients also completed the test at baseline.

The mean baseline digit symbol coding scores for each tumour type groups were as follows: 65.7 (SD 14.5) for the low-grade glioma patients; 53.8 (SD 23.3) for the high-grade glioma patients; 58.5 (SD 21.5) for the metastasis group; 47.7 (SE 23.7) for the meningioma patients and 62.1 (SD 25.5) for the 'other' tumour group.

There was a significant main effect of the covariates age, $F(1,97) = 32.250$, $p < 0.001$, partial $\eta^2 = 0.250$; and NART score, $F(1,97) = 27.236$, $p < 0.001$, partial $\eta^2 = 0.219$, in the model. Older patients were significantly more likely to score lower on this test and patients with higher (poorer) NART scores were also more likely to perform less well. The effect of sex was not significant, $F(1,97) = 0.055$, $p = 0.816$, partial $\eta^2 = 0.001$. The effect of tumour type was not significant in the model that included the effects of age and NART score, $F(4,97) = 0.549$, $p = 0.700$, partial $\eta^2 = 0.022$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.6, on digit symbol coding at baseline for each tumour type are as follows: 60.1 (SE 4.3) for the low-grade glioma group; 54.7 (SE 2.6) for the high-grade glioma group; 59.8 (SE 5.0) for the metastasis group; 53.8 (SE 9.4) for the meningioma group and 60.6 (SE 5.8) for the 'other' tumour group. LSD pairwise comparisons showed there to be no significant difference between the low-grade and high-grade glioma groups on digit symbol coding.

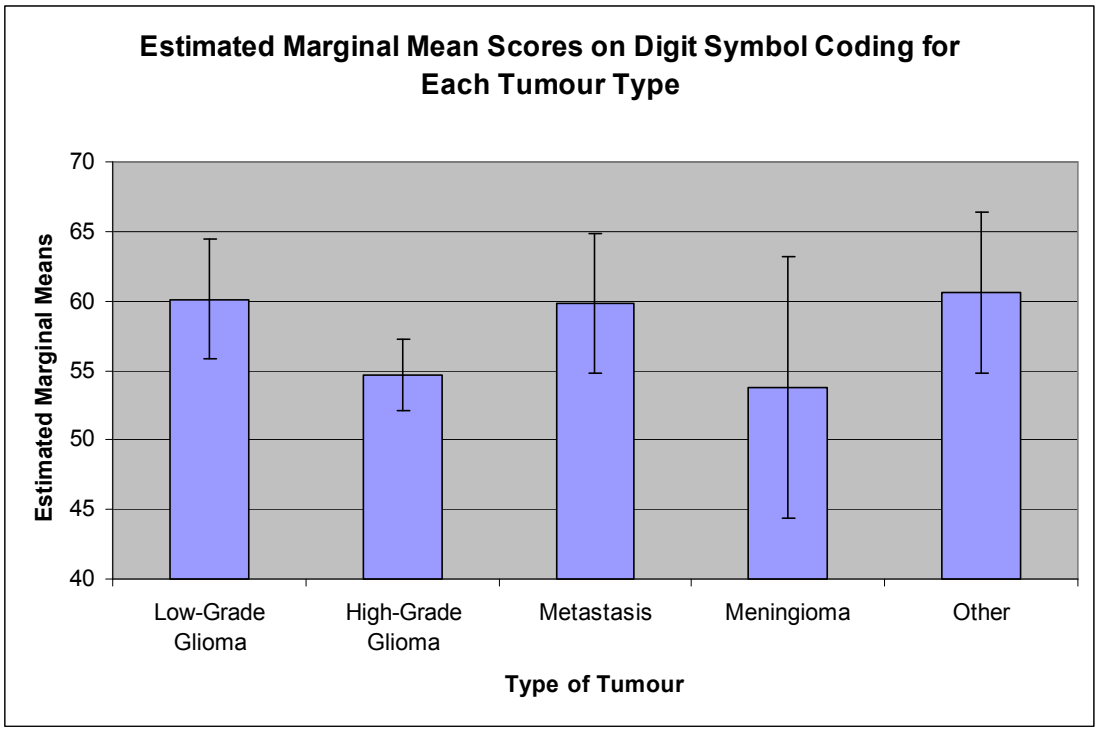


Figure 8.6. Baseline digit symbol coding test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.2.7 Letter-Number Sequencing

Nineteen low-grade glioma patients completed this test at baseline. Thirty-nine high-grade glioma patients, 9 metastasis patients, 10 meningioma patients and 10 patients with 'other' tumours also completed the letter-number sequencing test at baseline.

The mean score for the low-grade glioma group was 10.5 (SD 2.8) and for the high-grade glioma group was 8.6 (SD 3.6). The mean score for the metastasis group was 9.3 (SD 3.0), for the meningioma group was 10.3 (SD 4.1) and for the 'other' tumour group was 9.7 (SD 2.7).

The covariate age had a significant main effect in the model, $F(1,75) = 12.627$, $p = 0.001$, partial $\eta^2 = 0.144$. Younger patients were significantly more likely to have better scores on this test. The effect of the covariate NART was also significant, $F(1,75) = 24.071$, $p < 0.001$, partial $\eta^2 = 0.243$. Lower (better) NART scores were significantly associated with better letter-number sequencing scores. Sex had no

significant main effect in the model, $F(1,75) = 0.096$, $p = 0.758$, partial $\eta^2 = 0.001$. There was no significant main effect of tumour type in the model that included the effects of age and NART score, $F(4,75) = 0.842$, $p = 0.503$, partial $\eta^2 = 0.043$.

The estimated marginal mean scores for each group, adjusted for age and NART score and shown in Figure 8.7, on the letter-number sequencing test at baseline were as follows: 9.9 (SE 0.7) for the low-grade glioma group; 8.8 (SE 0.5) for the high-grade glioma group; 8.3 (SE 1.1) for the metastasis group; 10.8 (SE 1.6) for the meningioma group; and 10.1 (SE 0.9) for the 'other' tumour group. Pairwise comparisons of the low-grade and high-grade glioma groups using LSD tests revealed no significant difference between the performance of the two groups.

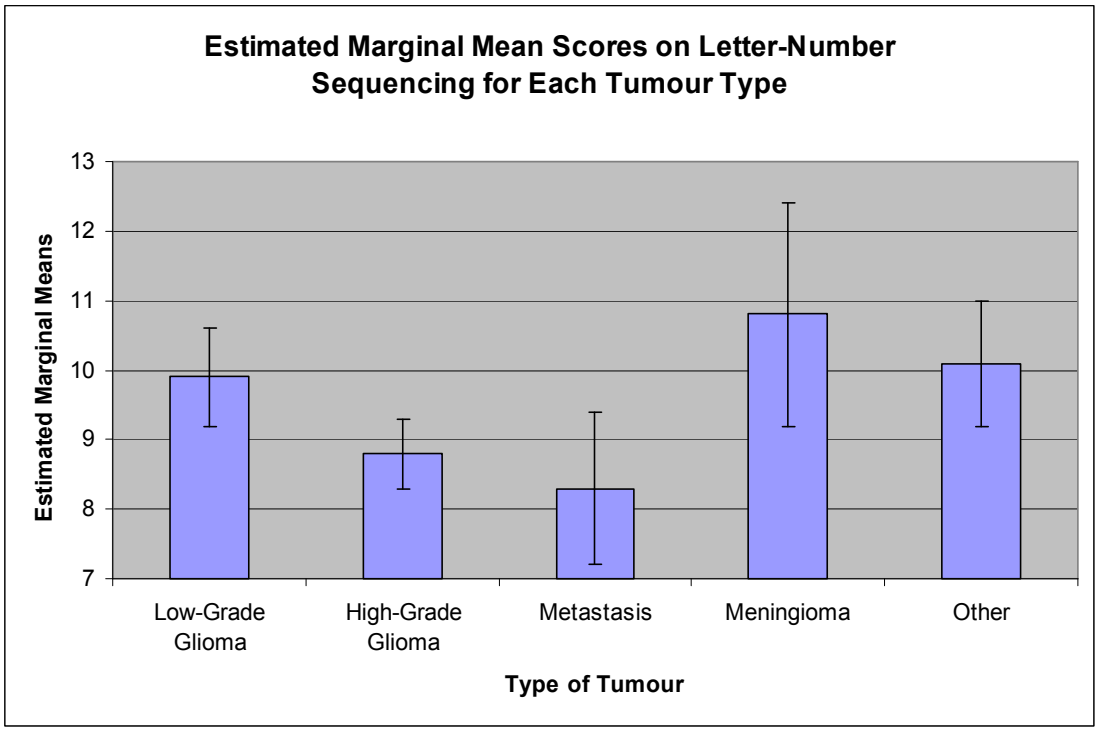


Figure 8.7. Baseline letter-number sequencing test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.8 Williams Delayed Recall Test (EFIT)

Twenty-two patients with a low-grade glioma and 49 with a high-grade glioma completed this test at baseline. A further 14 patients with a metastasis, 17 with a meningioma and 11 with 'other' tumour types also completed the Williams Delayed Recall Test at baseline.

The mean score for the low-grade glioma group was 7.6 (SD4.4), for the high-grade glioma group was 8.9 (SD 4.9) and for the metastasis group was 8.1 (SD 5.6). The meningioma group had a mean score of 7.4 (SD 6.4) and the 'other' tumour group mean score was 7.9 (SD 4.3).

There was a significant main effect of both of the covariates age, $F(1,101) = 11.163$, $p = 0.001$, partial $\eta^2 = 0.100$; and NART score, $F(1,101) = 4.584$, $p = 0.035$, partial $\eta^2 = 0.043$, in the model. Older age and higher (poorer) NART scores were significantly associated with higher (poorer) scores on this test. Sex had no significant effect in the model, $F(1,101) = 1.172$, $p = 0.281$, partial $\eta^2 = 0.011$. Tumour type had no significant main effect in the model that included the effects of age and NART score, $F(4,101) = 0.253$, $p = 0.907$, partial $\eta^2 = 0.010$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.8, for each tumour type group on the Williams Delayed Recall Test at baseline, were 8.5 (SE 1.2) for the low-grade glioma group; 8.7 (SE 0.7) for the high-grade glioma group; 7.7 (SE 1.3) for the metastasis group; 7.2 (SE 2.6) for the meningioma group and 7.7 (SE 1.5) for the 'other' tumour group. LSD pairwise comparisons of the low-grade and high-grade glioma groups revealed no significant difference in the performance of the two groups on the Williams Delayed Recall Test.

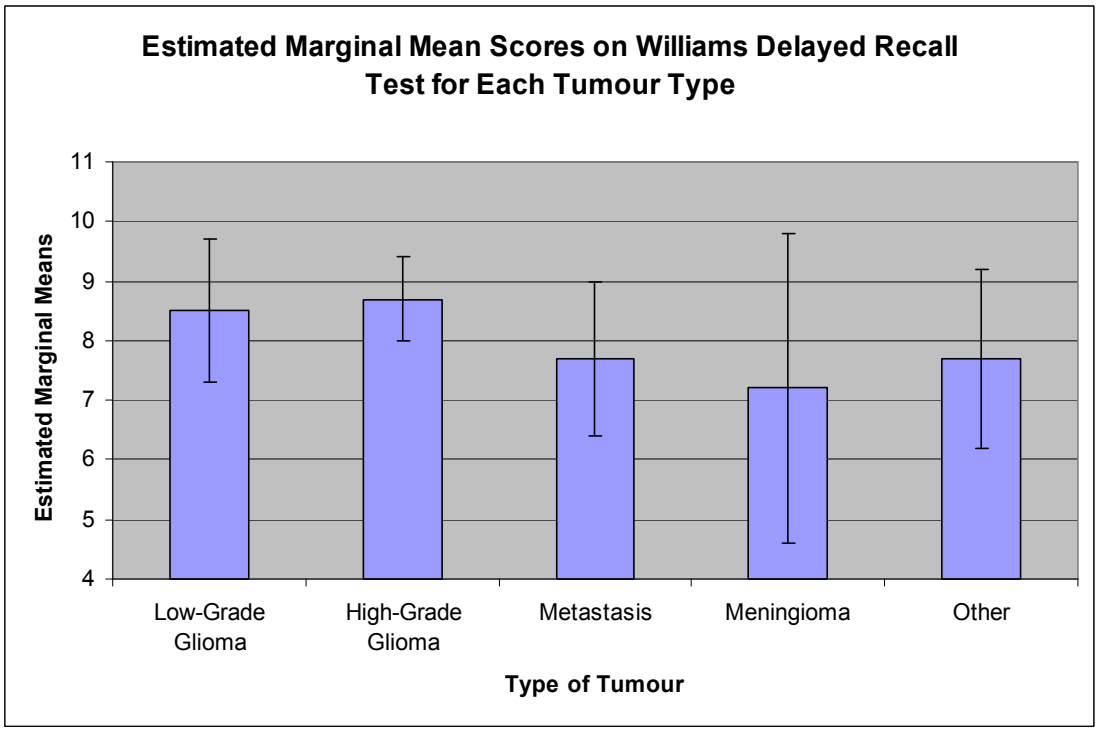


Figure 8.8. Baseline Williams Delayed Recall Test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.9 Nine Hole Peg Test (Right Hand, EFIT)

Twenty-two low-grade glioma patients, 49 high-grade glioma patients, 14 metastasis patients, 17 meningioma patients and 11 'other' tumour patients completed this measure at baseline.

The mean score for the low-grade group was 12.8 (SD 2.0) and for the high-grade group was 15.3 (SD 5.6). The metastasis group scored a mean of 13.5 (SD 2.2), the meningioma group scored a mean of 15.8 (SD 4.8) and the 'other' tumour group mean was 15.4 (SD 4.0).

There was a significant main effect of the covariate age in the model, $F(1,101) = 8.191$, $p = 0.005$, partial $\eta^2 = 0.075$. Younger patients tended to perform faster on this test. The covariate NART also had a significant main effect, $F(1,101) = 10.372$, $p = 0.002$, partial $\eta^2 = 0.093$. Patients with lower (better) NART scores also tended to

perform more quickly. Sex had no significant main effect in the model, $F(1,101) = 0.497$, $p = 0.483$, partial $\eta^2 = 0.005$. The effect of tumour type was not significant in the model that included the effects of age and NART score, $F(4,101) = 0.992$, $p = 0.416$, partial $\eta^2 = 0.038$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.9, for each of the tumour group on the right-hand nine hole peg test were 13.4 (SE 1.0) for the low-grade glioma group; 15.1 (SE 0.6) for the high-grade glioma group; 13.5 (SE 1.2) for the metastasis group; 15.4 (SE 2.2) for the meningioma group and 15.6 (SE 1.3) for the 'other' tumour group. Pairwise comparisons using LSD tests showed no significant difference between the low-grade and high-grade glioma groups.

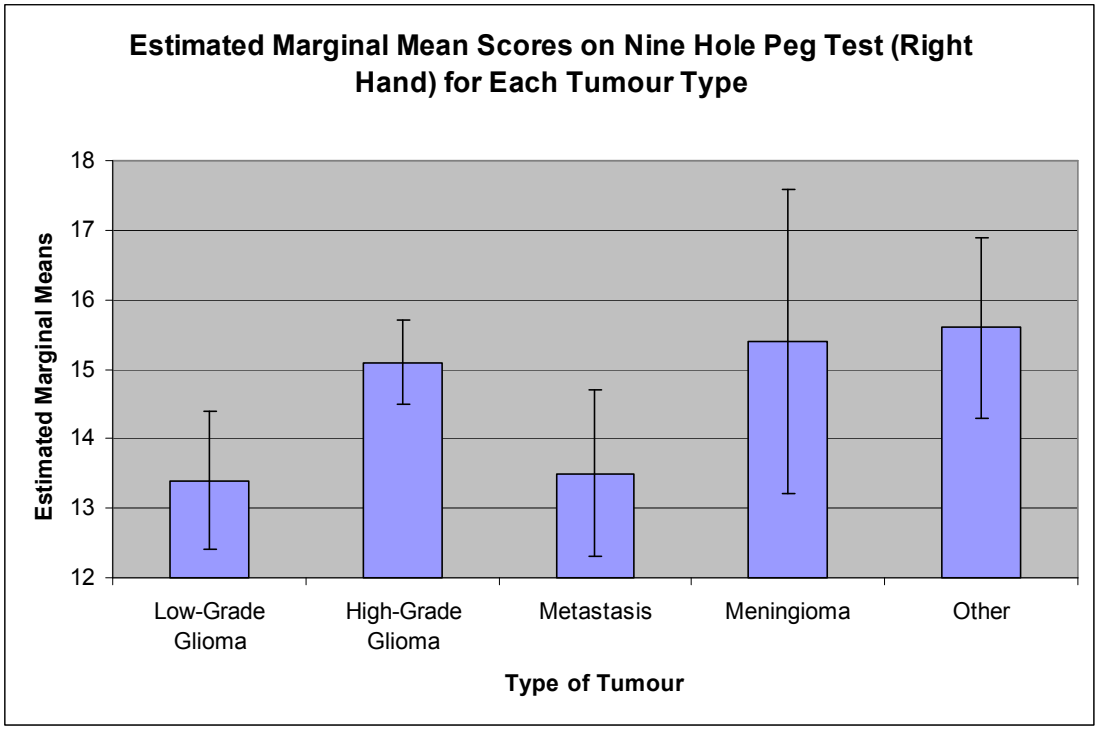


Figure 8.9. Baseline right hand nine hole peg test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.10 Nine Hole Peg Test (Left Hand, EFIT)

Twenty-two low-grade glioma patients, 46 high-grade glioma patients, 14 metastasis patients, 17 meningioma patients and 11 patients with 'other' tumours all completed this test at baseline.

The mean score for the low-grade glioma group was 13.3 (SD 1.9) and for the high-grade glioma group was 15.5 (SD 3.4). The metastasis group had a mean score of 15.4 (SD 4.1), the meningioma group had a mean of 17.5 (SD 4.0) and the 'other' tumour group had a mean score of 15.4 (SD 3.1).

There was a significant main effect of the covariate age in the model, $F(1,98) = 9.759$, $p = 0.002$, partial $\eta^2 = 0.091$. Older patients were significantly more likely to take longer to complete the test. The effect of the covariate NART was not significant, $F(1,98) = 1.585$, $p = 0.211$, partial $\eta^2 = 0.016$. Sex had no significant main effect in the model, $F(1,98) = 1.227$, $p = 0.271$, partial $\eta^2 = 0.012$. The effect of tumour type was not significant in the model that included the effects of age and NART score, $F(4,98) = 1.416$, $p = 0.234$, partial $\eta^2 = 0.055$.

The estimated marginal mean scores for the nine hole peg test (left hand), adjusted for age and NART score and shown in Figure 8.10, for each tumour group, were 13.7 (SE 0.8) for the low-grade glioma group; 15.3 (SE 0.5) for the high-grade glioma group; 15.2 (SE 0.9) for the metastasis group; 16.8 (SE 1.7) for the meningioma group and 15.9 (SE 1.0) for the 'other' tumour group. Pairwise comparisons using LSD tests revealed no significant difference between the low-grade and high-grade glioma groups.

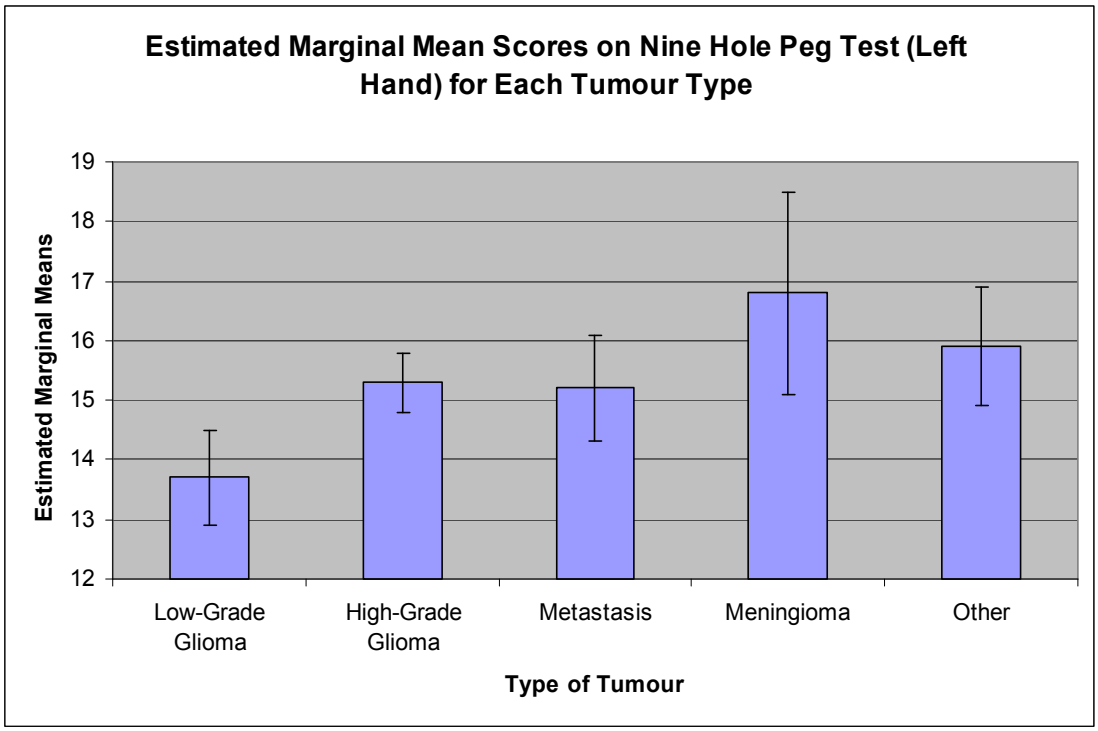


Figure 8.10. Baseline left hand nine hole peg test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.11 Timed Ten Metre Walk (EFIT)

There were 21 low-grade glioma patients, 44 high-grade glioma patients, 12 metastasis patients, 12 meningioma patients and 11 patients with 'other' tumours who completed the timed ten metre walk at baseline.

The mean score for the low-grade glioma group was 6.1 (SD 1.0) and for the high-grade glioma group was 6.8 (SD 1.4). The metastasis group had a mean score of 8.2 (SD 2.8), the meningioma group scored a mean of 7.9 (SD 1.6) and the 'other' tumour group scored a mean of 6.8 (SD 2.4).

There was a significant main effect in the model of both covariates age, $F(1,88) = 5.514$, $p = 0.021$, partial $\eta^2 = 0.059$; and NART score, $F(1,88) = 4.689$, $p = 0.033$, partial $\eta^2 = 0.051$. Older patients and those patients with higher (poorer) NART scores were significantly more likely to take longer to complete the test. The effect of sex was not significant, $F(1,88) = 0.449$, $p = 0.505$, partial $\eta^2 = 0.005$. The effect of

tumour type was not significant in the model that included the effects of the covariates, $F(4,88) = 2.066$, $p = 0.092$, partial $\eta^2 = 0.086$.

The estimated marginal mean scores on the timed ten metre walk, adjusted for age and NART score and shown in Figure 8.11, for each tumour group, were as follows: 6.4 (SE 0.4) for the low-grade glioma group; 6.7 (SE 0.3) for the high-grade glioma group; 8.1 (SE 0.5) for the metastasis group; 7.2 (SE 0.9) for the meningioma group and 6.8 (SE 0.5) for the 'other' tumour group. Pairwise comparisons using LSD test showed no significant difference between the performance of the low-grade and high-grade glioma groups.

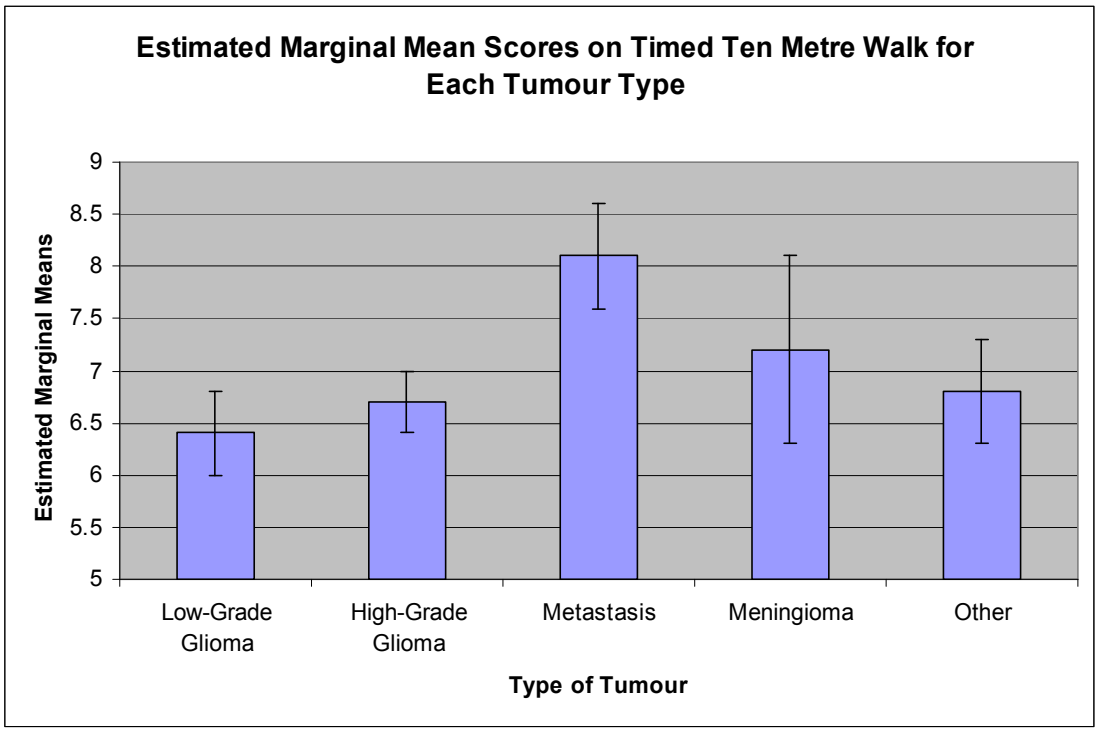


Figure 8.11. Baseline timed ten metre walk test scores for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.12 Hospital Anxiety and Depression Scale

There were 23 low-grade glioma patients, 49 high-grade glioma patients, 14 metastasis patients, 16 meningioma patients and 11 patients with 'other' tumours who completed the Hospital Anxiety and Depression Scale.

8.2.12.1 Anxiety Scores

The mean anxiety scores from the scale for each tumour type group were as follows: 6.5 (SD 3.6) for the low-grade glioma group; 7.9 (SD 4.4) for the high-grade glioma group; 8.9 (SD 4.5) for the metastasis group; 9.3 (SD 4.3) for the meningioma group and 9.4 (SD 5.2) for the 'other' tumour group.

There was no significant main effect in the model of either of the covariates age, $F(1,101) = 0.360$, $p = 0.550$, partial $\eta^2 = 0.004$; or NART score, $F(1,101) = 0.339$, $p = 0.562$, partial $\eta^2 = 0.003$. The effect of sex was also not significant, $F(1,101) = 0.209$, $p = 0.648$, partial $\eta^2 = 0.002$. The effect of tumour type was not significant in the model that included the effects of age and NART score, $F(4,101) = 1.174$, $p = 0.327$, partial $\eta^2 = 0.044$.

The estimated marginal mean anxiety scores, adjusted for age and NART score and shown in Figure 8.12, were 6.2 (SE 1.0) for the low-grade glioma group, 7.9 (SE 0.6) for the high-grade glioma group, 9.0 (SE 1.2) for the metastasis group, 8.7 (SE 2.4) for the meningioma group and 9.2 (SD 1.4) for the 'other' tumour group. LSD tests showed no significant difference in anxiety levels reported by the low-grade and high-grade glioma groups.

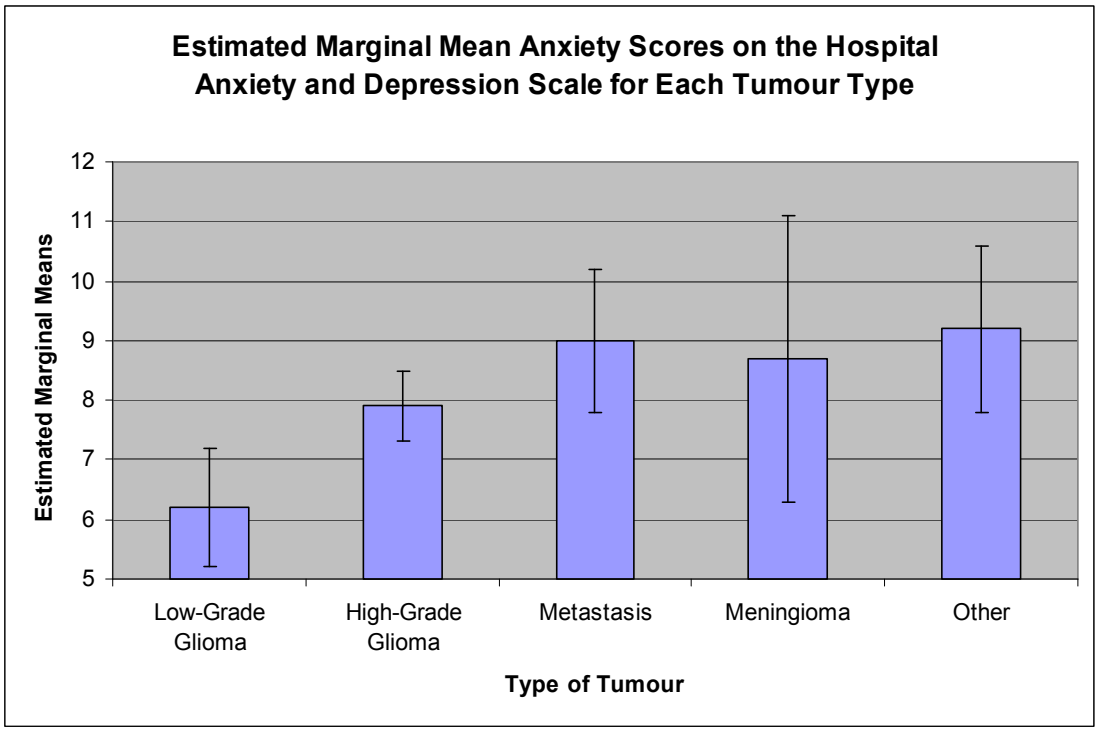


Figure 8.12. Baseline anxiety scores on the Hospital Anxiety and Depression Scale for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.12.2 Depression Scores

The mean depression scores on the Hospital Anxiety and Depression Scale were 3.5 (SD 3.8) for the low-grade glioma patients, 4.5 (SD 4.3) for the high-grade glioma patients, 5.0 (SD 2.9) for the metastasis group, 5.9 (SD 4.5) for the meningioma group and 5.8 (SD 4.0) for the 'other' tumour group.

There was no significant main effect in the model of either of the covariates age, $F(1,101) = 0.705$, $p = 0.403$, partial $\eta^2 = 0.007$; or NART score, $F(1,101) = 1.962$, $p = 0.164$, partial $\eta^2 = 0.019$. The effect of sex was not significant in the model, $F(1,101) = 2.586$, $p = 0.111$, partial $\eta^2 = 0.0.25$. The effect of tumour type was not significant in the model that included the effects of age and NART score, $F(4,101) = 1.064$, $p = 0.378$, partial $\eta^2 = 0.040$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.13, on the depression scale of the Hospital Anxiety and Depression Scale,

were as follows: 3.2 (SE 0.9) for the low-grade glioma patients; 4.6 (SE 0.6) for the high-grade glioma patients; 5.4 (SE 1.1) for the metastasis patients; 2.4 (SE 2.2) for the meningioma patients and 5.5 (SE 1.3) for the 'other' tumour patients. Pairwise comparisons using LSD tests revealed no significant differences between the levels of depression reported by the low-grade and high-grade glioma groups, as measured by the Hospital Anxiety and Depression Scale.

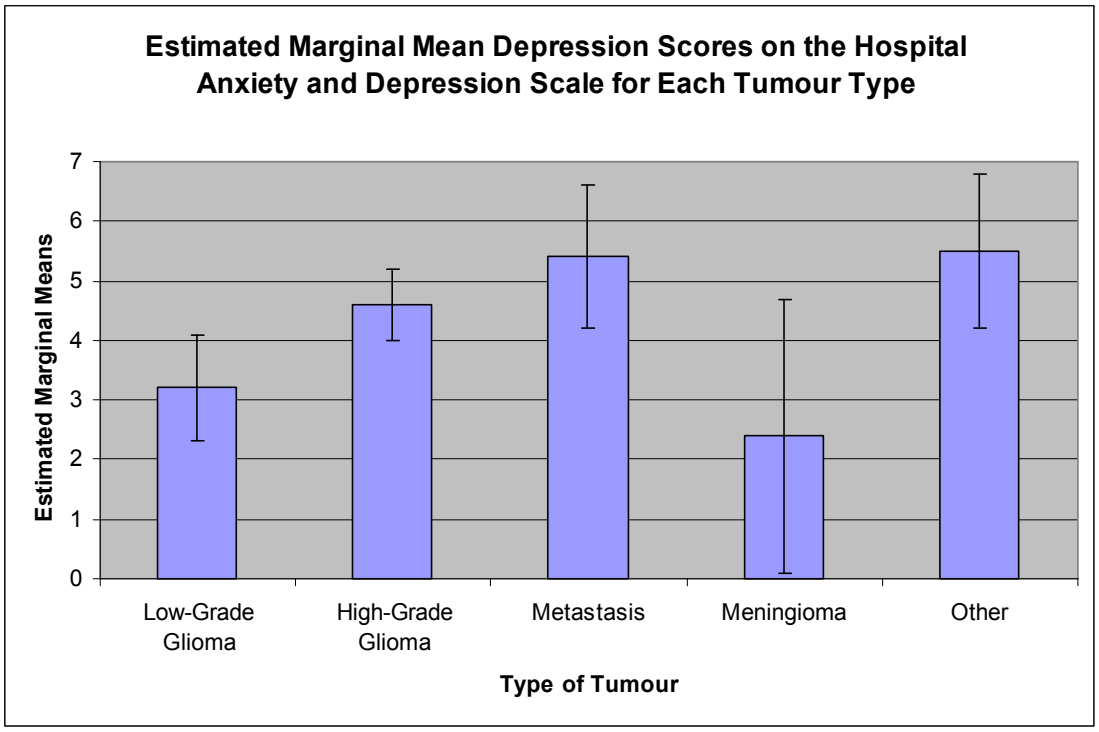


Figure 8.13. Baseline depression scores on the Hospital Anxiety and Depression Scale for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.2.12.3 Hospital Anxiety and Depression Scale – Total Score

The mean total score on the Hospital Anxiety and Depression Scale for the low-grade glioma group was 10.0 (SD 6.6), for the high-grade glioma group was 12.4 (SD 7.9), for the metastasis group was 13.9 (SD 6.3), for the meningioma group was 15.3 (SD 8.2) and for the 'other' tumour group was 15.2 (SD 7.6).

The covariate age had no significant main effect in the model, $F(1,101) = 0.645$, $p = 0.424$, partial $\eta^2 = 0.006$; nor did the covariate NART score, $F(1,101) = 1.197$, $p = 0.277$, partial $\eta^2 = 0.012$. Sex had no significant effect in the model, $F(1,101) = 1.281$, $p = 0.260$, partial $\eta^2 = 0.013$. The effect of tumour type was not significant tin the model that included the effects of age and NART score, $F(1,101) = 1.270$, $p = 0.287$, partial $\eta^2 = 0.048$.

The estimated marginal mean total scores, adjusted for age and NART score and shown in Figure 8.14, were 9.4 (SE 1.7) for the low-grade glioma group, 12.5 (SE 1.1) for the high-grade glioma group, 14.4 (SE 2.1) for the metastasis group, 11.1 (SE 4.0) for the meningioma group and 14.7 (SE 2.4) for the 'other' tumour group. LSD pairwise comparisons revealed no significant difference between the low-grade and high-grade glioma groups.

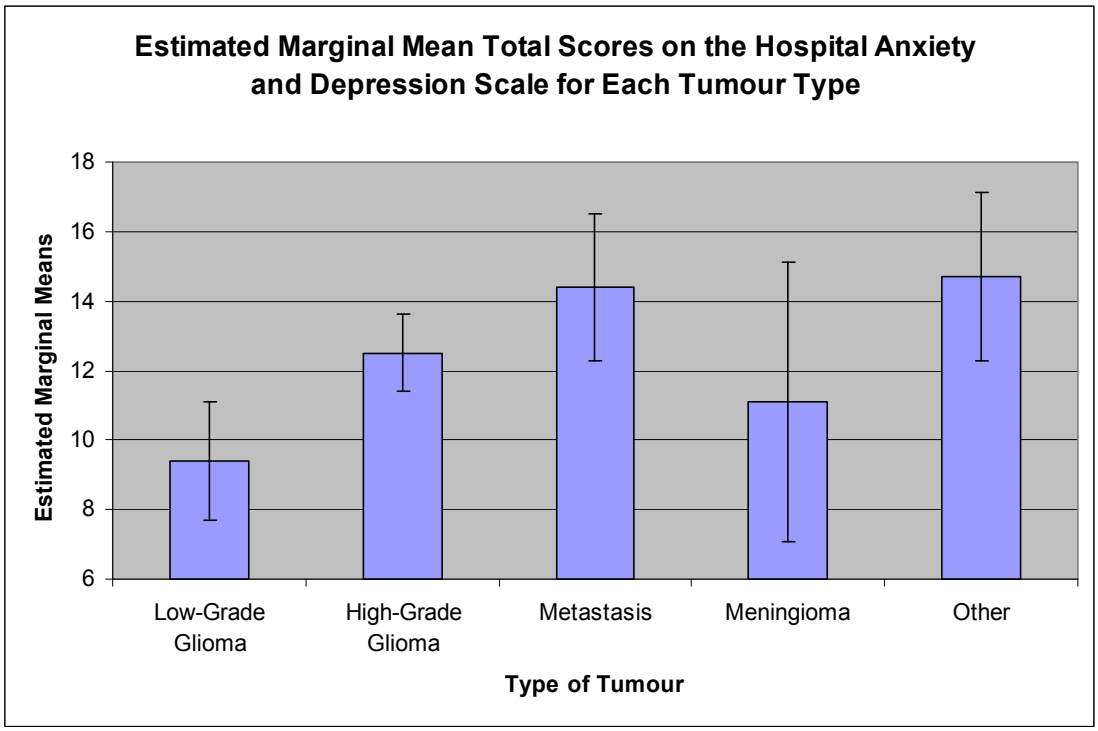


Figure 8.14. Baseline total scores on the Hospital Anxiety and Depression scales for each histological tumour type group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means

Table 8.1. Comparison of histological groups on cognitive and mood scales at baseline.

Test	Effect of tumour type		
	F	Sig. (p)	n^{2}*
Inspection Time (All Data)	1.280	0.283	0.050
Inspection Time (Valid Data)	0.683	0.605	0.032
Rey Auditory Verbal Learning Test (Total)	2.259	0.071	0.110
Trail Making Test Part B (secs)	0.830	0.510	0.039
Verbal Fluency (Total)	5.700	< 0.001	0.228
Digit Symbol Coding (Total)	0.549	0.700	0.022
Letter-Number Sequencing (Total)	0.842	0.503	0.043
EFIT Williams Delayed Recall Test (total)	0.253	0.907	0.010
EFIT Nine Hole Peg Test (Right Hand, secs)	0.992	0.416	0.038
EFIT Nine Hole Peg Test (Left Hand, secs)	1.416	0.234	0.055
EFIT Timed Ten Metre Walk (secs)	2.066	0.092	0.086
Hospital Anxiety and Depression Scale –Anxiety Score	1.174	0.327	0.044
Hospital Anxiety and Depression Scale – Depression Score	1.064	0.378	0.040
Hospital Anxiety and Depression Scale – Total Score	1.270	0.287	0.048

*n^{2} = the proportion of variance accounted for by tumour type. HGG = high-grade glioma; LGG = low-grade glioma

8.3 Tumour Location – Lobe

The brain tumour cohort were classified into the following seven main subtypes, according to the location of their tumour, as stated by formal neuro-radiology reports:

(i) Frontal lobe
(ii) Temporal lobe
(iii) Limbic
(iv) Parietal lobe
(v) Occipital Lobe
(vi) Multiple Lobes
(vii) Other

Of the total cohort of brain tumour patients (n = 118), 34 patients had a tumour located in the frontal lobes, 16 had a temporal lobe tumour, 8 had a tumour located within the limbic system, 19 had a parietal tumour, 5 had an occipital tumour and 23 had tumour that invaded more than one lobe ('multiple lobes'). The remaining 13 had a tumour located in a different region in the brain ('other') and this included those patients with a pituitary tumour.

8.3.1 Inspection Time Scores: All Inspection Time Data

When all inspection time data was analysed, the group of patients with a frontal tumour scored a mean of 118.2 (SD 20.3). The temporal tumour group scored a mean of 122.0 (SD 11.9) and the limbic group had a mean score of 120.7 (SD10.8) at baseline. The patients with a parietal tumour scored 109.9 (SD 18.7) and the patients with an occipital lobe tumour scored a mean of 82.4 (SD 24.0). The 'multiple lobe' tumour group had a mean score of 109.9 (SD 21.2) and the patients with tumour located elsewhere ('other') scored 116.7 (SD 16.4) on the inspection time measure at baseline testing.

In this model, there was a significant main effect of the covariates age, $F(1,96) = 39.264$, $p < 0.001$, partial $\eta^2 = 0.290$; and NART score, $F(1,96) = 8.645$, $p = 0.004$, partial $\eta^2 = 0.083$, were significant. Older patients and those with higher (poorer) NART scores had significantly lower inspection time scores. The effect of sex was not significant, $F(1,96) = 0.690$, $p = 0.408$, partial $\eta^2 = 0.007$. Tumour lobe had a significant effect on baseline inspection time scores, in the model that included the effects of age and NART score, $F(6,96) = 3.732$, $p = 0.002$, partial $\eta^2 = 0.189$.

The estimated marginal mean inspection time scores, adjusted for age and NART score and shown in Figure 8.15, for each tumour lobe group were118.1 (SE 2.6) for the frontal lobe group; 121.0 (SE 4.0) for the temporal lobe group; 112.1 (SE 6.0) for the limbic group; 113.3 (SE 3.7) for the parietal group; 87.7 (SE 7.1) for the occipital group; 109.5 (SE 3.5) for the group with more than one lobe affected and 119.4 (SE 4.3) for the group with tumours located elsewhere ('other'). Pairwise comparisons using LSD tests revealed that the occipital lobe tumour group performed significantly worse than each of the other tumour lobe groups ($p < 0.001$ for the frontal group; $p = 0.001$ for the temporal lobe group; $p = 0.011$ for the limbic group; $p = 0.002$ for the parietal group; $p = 0.007$ for the multiple lobe group and $p < 0.001$ for the 'other' group). The frontal lobe group showed a tendency towards better performance by comparison with the multiple lobes tumour group but this difference did not reach the conventional level of statistical significance ($p = 0.054$).

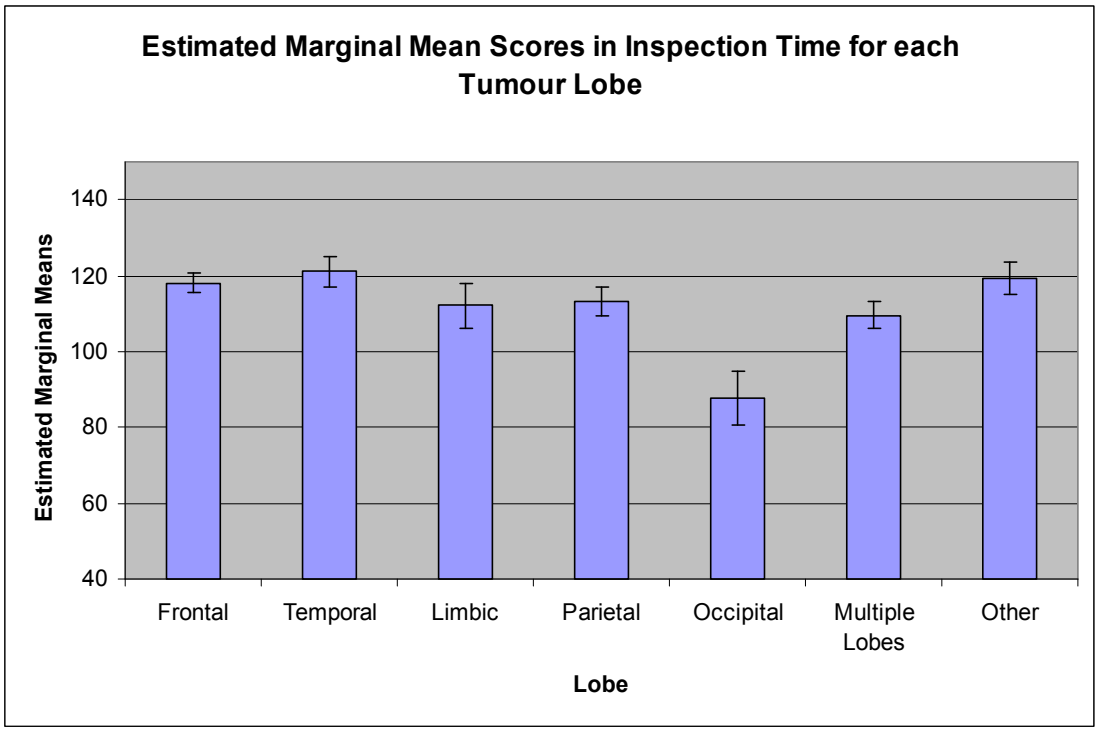


Figure 8.15. Baseline inspection time scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.3.2 Inspection Time Scores: Valid Inspection Time Data

When only 'valid' inspection time scores were entered into the model, this resulted in the 29 patients in the frontal tumour group, 15 in the temporal lobe group, 7 in the limbic group, 16 in the parietal group, 1 in the occipital group, 16 in the multiple lobes group and 12 patients with tumours located in 'other' areas.

The mean baseline inspection time score for the frontal tumour group who met validity criteria was 124.7 (SD 12.5). The temporal group had a mean of 122.0 (SD 11.9) and the limbic group had a mean of 120.7 (SD 10.8). The parietal group had a mean score of 113.7 (SD 15.8), the single patient with valid data in the occipital group had a score of 100.0 (SD.0.0) and the multiple lobe group had a mean of 118.5 (SD 12.7). The group of patients with tumours located elsewhere ('other') also had a mean score of 118.5 (SD 15.8).

The covariate age had a significant main effect in this model, $F(1,81) = 21.829$, $p < 0.001$, partial $\eta^2 = 0.212$. Older patients were more likely to have poorer inspection time scores. Sex also had a significant effect on valid baseline inspection time scores, $F(1,81) = 5.180$, $p = 0.025$, partial $\eta^2 = 0.060$. Female patients were significantly more likely to have poorer inspection time scores than male participants. NART had no significant effect in the model, $F(1,81) = 2.842$, $p = 0.096$, partial $\eta^2 = 0.034$. Tumour lobe also had no significant overall effect in the model that included the effects of age and NART score, $F(6,81) = 1.268$, $p = 0.281$, partial $\eta^2 = 0.086$.

The estimated marginal mean valid inspection time scores, adjusted for age and NART score and shown in Figure 8.16, for each tumour lobe group were 124.3 (SE 2.1) for the frontal lobe tumour group; 122.1 (SE 3.0) for the temporal lobe group; 116.4 (SE 4.5) for the limbic group, 116.0 (SE 3.0) for the parietal group; 112.6 (SE 11.8) for the occipital group; 117.6 (SE 3.0) for the group with multiple lobe involvement and 120.5 (SE 3.4) for the 'other' lobe group.

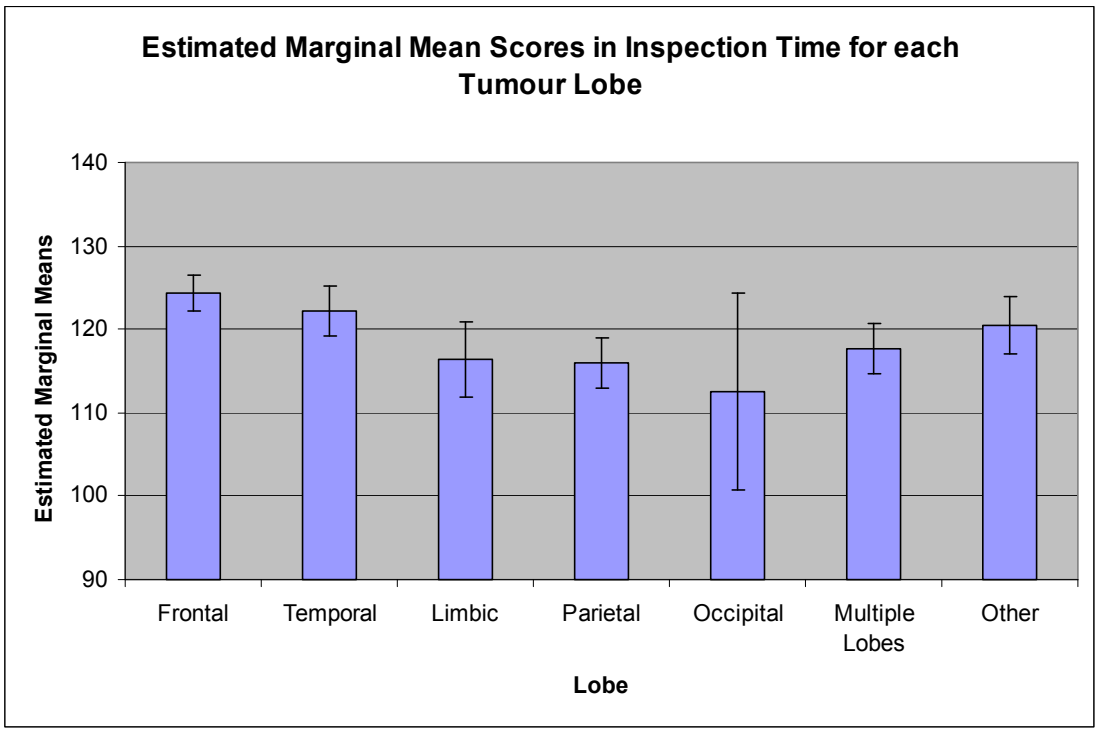


Figure 8.16. Baseline inspection time scores including only valid scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.3.3 Rey Auditory Verbal Learning Test

Twenty-five frontal lobe tumour patients completed this test at baseline. There were 13 patients in the temporal lobe group, 5 in the limbic group, 12 parietal lobe patients, 4 occipital lobe patients, 17 'multiple lobe' tumour patients and 11 patients in the 'other' location group who also completed the Rey Auditory Verbal Learning Test at baseline.

The mean score for the frontal lobe group on this measure at baseline was 64.5 (SD 17.7) and for the temporal lobe group was 57.9 (SD 19.8). The limbic group mean score was 53.0 (SD 14.0), the parietal lobe group had a mean score of 64.6 (SD 21.3), the occipital lobe group had a mean score of 46.3 and the multiple lobe tumour group had a mean of 59.5 (SD 14.9). The group of patients with tumours located in 'other' areas scored a mean of 69.5 (SD 14.0).

There was a significant main effect of the covariate age on test scores in this model, $F(1,71) = 25.646$, $p < 0.001$, partial $\eta^2 = 0.265$. Younger patients were significantly more likely to have higher (better) scores. The covariate NART also had a significant main effect in the model, $F(1,71) = 19.447$, $p < 0.001$, partial $\eta^2 = 0.215$. Better NART scores were significantly associated with better RAVLT scores. The effect of sex did not reach the conventional level of statistical significance, $F(1,71) = 3.418$, $p = 0.069$, partial $\eta^2 = 0.046$. The effect of tumour location (lobe) was significant in the model that included the effects of age and NART score, $F(6,71) = 2.268$, $p = 0.046$, partial $\eta^2 = 0.161$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.17, for each tumour lobe group on the Rey Auditory Verbal Learning Test, were as follows: 62.2 (SE 2.9) for the frontal lobe group; 56.9 (SE 4.0) for the temporal lobe group; 47.6 (SE 8.0) for the limbic group; 62.8 (SE 4.4) for the parietal lobe group; 49.6 (SE 7.2) for the occipital lobe group; 61.4 (SE 3.5) for the multiple lobe group and 72.5 (SE 4.3) for the group of patients with tumours in 'other' brain

areas. Pairwise comparisons using LSD tests revealed that the 'other' tumour location group performed significantly better than the temporal lobe group (p = 0.010); the limbic group (p = 0.008); the occipital lobe group (p = 0.008) and the multiple lobe group (p = 0.050). The 'other' location group showed a tendency towards better performance than the frontal lobe group (p = 0.055). There were no other significant differences between the tumour lobe subgroups.

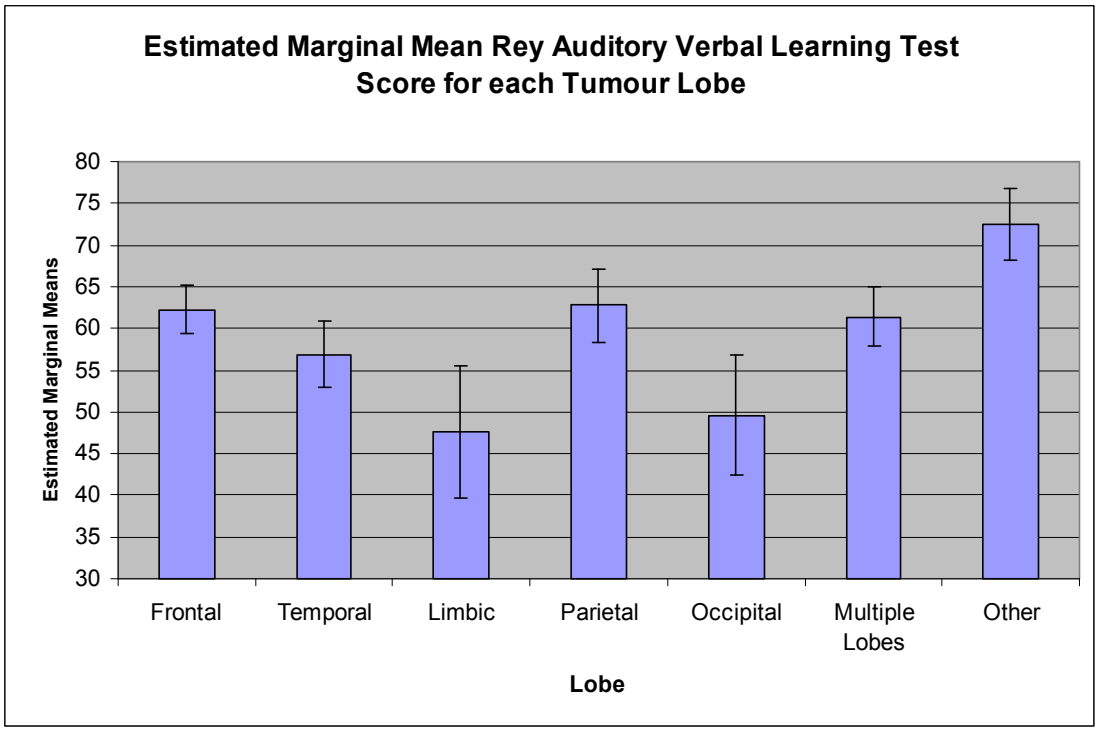


Figure 8.17. Baseline Rey Auditory Verbal Learning Test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.4 Trail Making Test Part B

Thirty patients with frontal tumours and 16 with temporal tumours completed the trail making test part B at baseline. Five patients in the limbic group, 12 in the parietal group, 3 with occipital tumours, 19 with tumours in 'multiple lobes' and 11 with tumours located elsewhere ('other') also completed the test.

The frontal tumour group scored a mean of 96.4 (SD 32.0) and the temporal lobe group had a mean score of 90.4 (SD 29.4). The limbic group had a mean of 99.0 (SD

20.1), the parietal group scored a mean of 103.2 (SD 39.1) and the mean score for the occipital lobe group was 105.9 (SD 67.4). The group of patients with a tumour invading multiple lobes scored a mean of 89.4 (SD 30.5) on the trail making test part B and the 'other' tumour location group had a mean score of 92.4 (SD 24.7).

There was a significant main effect of age in the model, $F(1,80) = 18.995$, $p < 0.001$, partial $\eta^2 = 0.192$. Older patients were more likely to take longer to complete the test. The covariate NART score also had a significant main effect in the model, $F(1,80) = 19.715$, $p < 0.001$, partial $\eta^2 = 0.198$. Those participants with higher (poorer) NART scores were significantly more likely to take longer to complete the trail making test part B. The effect of sex was not significant in the model, $F(1,80) = 3.305$, $p = 0.073$, partial $\eta^2 = 0.040$. Tumour lobe had no significant effect on test scores in the model that included the effects of the covariates, $F(6,80) = 0.846$, $p = 0.538$, partial $\eta^2 = 0.060$.

The estimated marginal mean scores for each tumour lobe group on the trail making test part B, adjusted for age and NART score and shown in Figure 8.18, were as follows: 97.3 (SE 5.0) for the frontal lobe group; 89.0 (SE 6.8) for the temporal lobe group; 107.7 (SE 12.5) for the limbic group; 101.1 (SE 7.9) for the parietal group; 111.5 (SE 16.7) for the occipital group; 90.2 (SE 6.3) for the multiple lobe group and 87.1 (SE 8.3) for the 'other' location group.

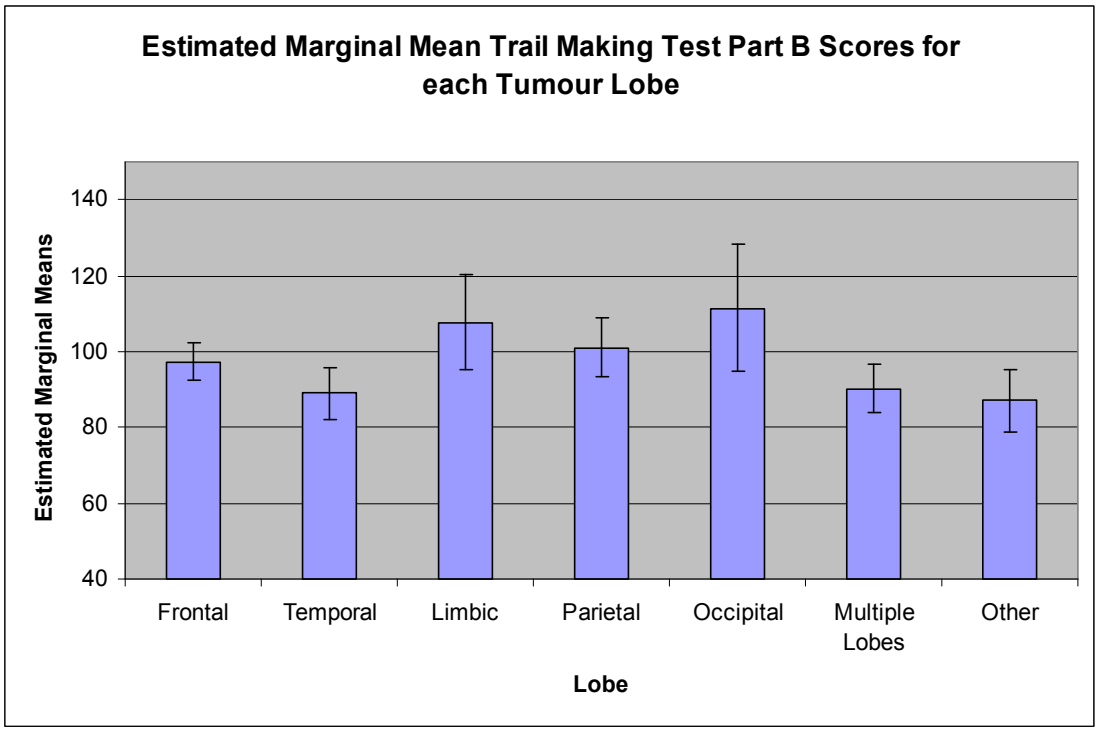


Figure 8.18. Baseline trail making test part b scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.5 Verbal Fluency

Twenty-eight patients in the frontal lobe group, 13 in the temporal lobe group, 7 with limbic tumours, 12 with parietal tumours, 4 with occipital lobe tumours, 18 with tumours in more than one lobe ('multiple lobes') and 9 with tumours elsewhere in the brain ('other') all completed the verbal fluency measure at baseline.

The mean score on the verbal fluency test for the frontal lobe group was 28.5 (SD 11.1) and for the temporal lobe group was 28.0 (SD 9.6). The limbic group had a mean score of 26.7 (SD 6.2), the parietal lobe group scored a mean of 34.5 (SD 18.4), the occipital lobe group mean was 28.8 (SD 5.1), the multiple lobe group scored a mean of 25.7 (SD 13.7) and the mean score for the 'other' location group was 29.9 (SD 9.9).

The covariate age had a significant main effect in the model, $F(1,75) = 10.782$, $p = 0.002$, partial $\eta^2 = 0.126$. Older participants tended to produce more words on the

verbal fluency measure. There was a significant main effect of the covariate NART score in the model, $F(1,75) = 36.864$, $p < 0.001$, partial $\eta^2 = 0.330$. Participants with higher (poorer) NART scores tended to perform less well on the verbal fluency measure. The effect of sex was not significant in the model, $F(1,75) = 0.430$, $p = 0.514$, partial $\eta^2 = 0.006$. Tumour lobe had no significant effect on test scores in the model that included the effects of the covariates, $F(6,75) = 1.124$, $p = 0.357$, partial $\eta^2 = 0.083$.

The estimated marginal mean verbal fluency scores, adjusted for age and NART score and shown in Figure 8.19, for each tumour lobe group were 27.4 (SE 1.9) for the frontal lobe group, 28.3 (SE 2.8) for the temporal lobe group, 23.4 (SE 3.8) for the limbic group, 34.1 (SE 3.3) for the parietal group, 29.7 (SE 5.0) for the occipital lobe group, 27.4 (SE 2.4) for the multiple lobes group and 32.6 (SE 3.3) for the 'other' location group.

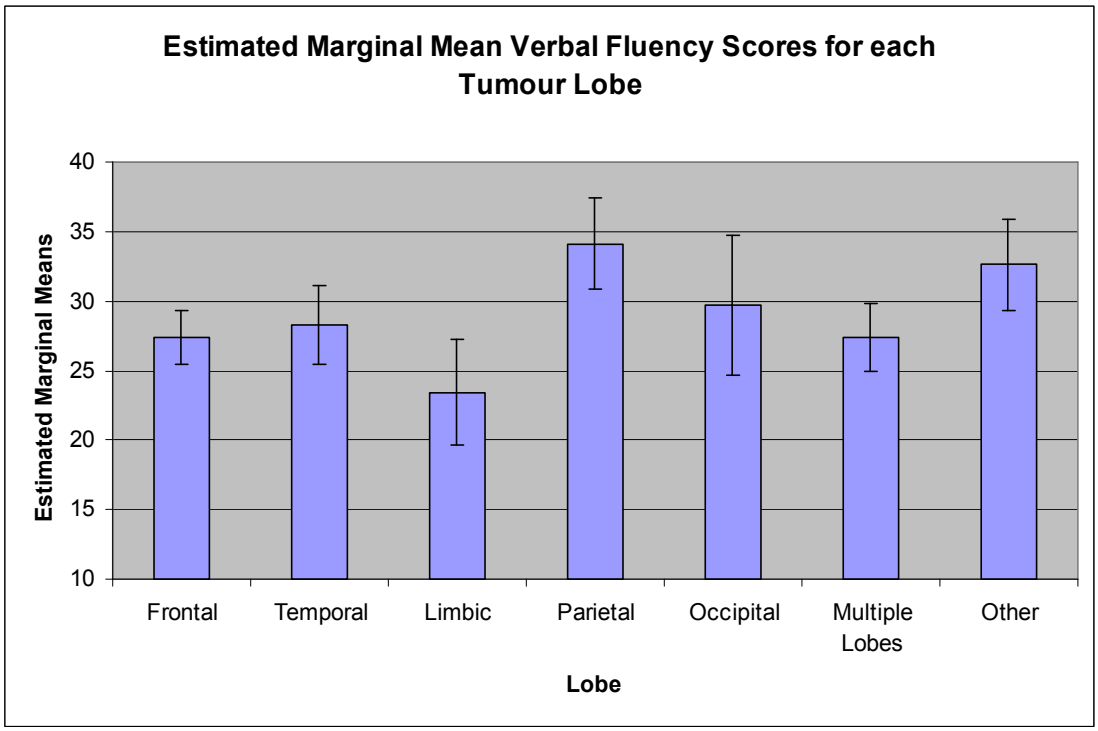


Figure 8.19. Baseline verbal fluency test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.6 Digit Symbol Coding

Thirty-four patients with frontal lobe tumours, 16 with temporal lobe tumours, 6 with limbic tumours, 17 with parietal lobe tumours, 4 with occipital lobe tumours, 21 with tumours in more than one lobe ('multiple lobes') and 13 with tumours located elsewhere within the brain ('other') all completed the digit symbol coding measure at baseline testing.

The mean score for the frontal lobe group was 59.3 (SD 20.7), for the temporal group was 65.2 (SD 20.8) and for the limbic group was 61.2 (SD 16.7). The parietal lobe group had a mean score of 47.5 (SD 19.8), the occipital lobe group mean was 37.5 (SD 24.8), the multiple lobe group mean score was 52.8 (SD 25.7) and the 'other' location group mean score was 58.6 (SD 22.5).

There was a significant main effect of the covariate age in the model, $F(1,95) = 42.733$, $p < 0.001$, partial $\eta^2 = 0.310$. Older participants were more likely to have lower (poorer) scores on the digit symbol coding measure. The covariate NART score also had a significant main effect in the model, $F(1,95) = 35.077$, $p < 0.001$, partial $\eta^2 = 0.270$. Those participants with higher (poorer) NART scores were more likely to have lower (poorer) digit symbol coding scores. Sex had no significant effect in the model, $F(1,95) = 0.014$, $p = 0.906$, partial $\eta^2 < 0.001$. The effect of tumour lobe approached the conventional level of statistical significance in the model that included the effects of the covariates age and NART score, $F(6,95) = 2.180$, $p = 0.052$, partial $\eta^2 = 0.121$.

The estimated marginal mean scores on the digit symbol coding test at baseline for each tumour lobe, adjusted for age and NART score and shown in Figure 8.20, were as follows: 58.3 (SE 2.9) for the frontal lobe group; 65.2 (SE 4.3) for the temporal lobe group; 50.2 (SE 7.1) for the limbic group; 47.5 (SE 4.2) for the parietal lobe group; 47.2 (SE 8.6) for the occipital lobe group; 53.3 (SE 3.7) for the multiple lobes group and 62.2 (SE 4.8) for the group of patients with tumours in 'other' areas of the brain. Since the main effect of tumour lobe approached the conventional level of significance, post-hoc analyses were carried out in this instance. Pairwise

comparisons, using LSD tests revealed that the frontal lobe group performed significantly better than the parietal lobe group ($p = 0.037$). The temporal lobe group and 'other' location group also performed significantly better than the parietal lobe group ($p = 0.004$ and $p = 0.023$, respectively). The temporal lobe group also performed significantly better than the multiple lobe group ($p = 0.039$) and showed a tendency towards better performance than the occipital lobe group ($p = 0.065$).The parietal lobe group performed significantly worse than the 'other' location group ($p = 0.023$).

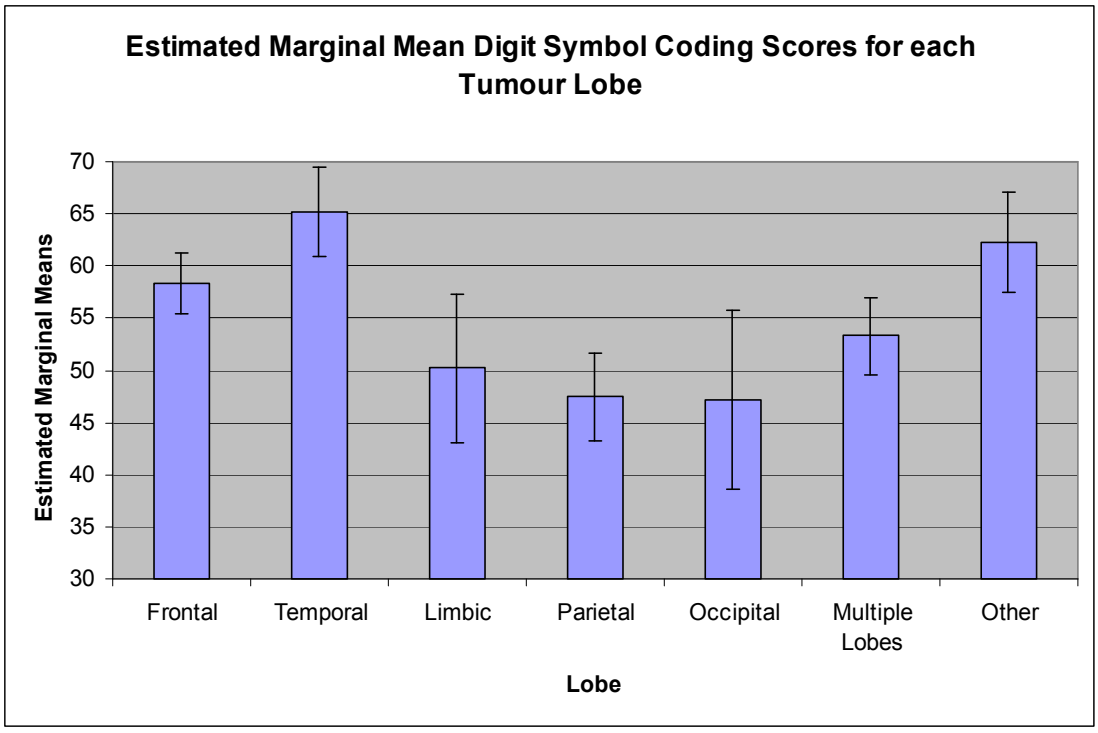


Figure 8.20. Baseline digit symbol coding test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.7 Letter-Number Sequencing

Twenty-seven members of the frontal lobe group, 12 in the temporal group, 6 with limbic tumours and 13 with parietal lobe tumours completed letter-number sequencing at baseline. Four patients with occipital lobe tumours, 18 with multiple lobe tumours, and 9 with tumours elsewhere ('other') also completed the test.

The mean score for the frontal lobe group was 10.2 (SE 3.2) and for the temporal lobe group was 10.2 (SE 2.8). The limbic group had a mean score of 8.7 (SE 3.4), the parietal lobe group mean was 10.0 (SE 3.6), the occipital lobe group scored a mean of 5.0 (SE 3.4), the multiple lobes group mean was 7.6 (SE 3.0) and the 'other' location group mean was 11.3 (SE 2.3).

There was a significant main effect of the covariate age on test scores in the model, $F(1,73) = 24.128$, $p < 0.001$, partial $\eta^2 = 0.248$. Older patients tended to perform less well on the letter-number sequencing test at baseline. There was also a significant main effect of the covariate NART score in the model, $F(1,73) = 31.643$, $p < 0.001$, partial $\eta^2 = 0.302$. Patients with lower (better) NART scores tended to have higher (better) scores on the test. There was no significant effect of sex in the model, $F(1,73) = 0.035$, $p = 0.852$, partial $\eta^2 < 0.001$. The effect of tumour lobe was significant in the model that included the effects of age and NART score, $F(6,73) = 5.166$, $p = 0.001$, partial $\eta^2 = 0.298$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.21, for each tumour location group on the letter-number sequencing test at baseline were as follows: 9.8 (SE 0.5) for the frontal lobe group; 10.2 (SE 0.7) for the temporal lobe group; 7.8 (SE 1.1) for the limbic group; 9.6 (SE 0.7) for the parietal lobe group; 5.6 (SE 1.2) for the occipital lobe group; 8.0 (SE 0.6) for the multiple lobes group and 12.2 (SE 0.8) for the group of patients with tumours located elsewhere in the brain ('other'). Pairwise comparisons, using LSD tests revealed that the 'other' location group performed significantly better than the frontal lobe group ($p = 0.018$); limbic group ($p = 0.002$); parietal lobe group ($p = 0.021$); occipital lobe group ($p < 0.001$) and the multiple lobe group ($p < 0.001$). The 'other' location group also showed a tendency towards better performance than the temporal lobe group, but this difference did not reach the conventional level of statistical significance ($p = 0.067$). The occipital lobe group performed significantly worse than the frontal lobe group ($p = 0.002$), temporal lobe group ($p = 0.002$) and parietal lobe group ($p = 0.005$). The frontal and temporal lobe groups also performed significantly better than the multiple lobe group ($p = 0.018$ and $p = 0.022$, respectively).

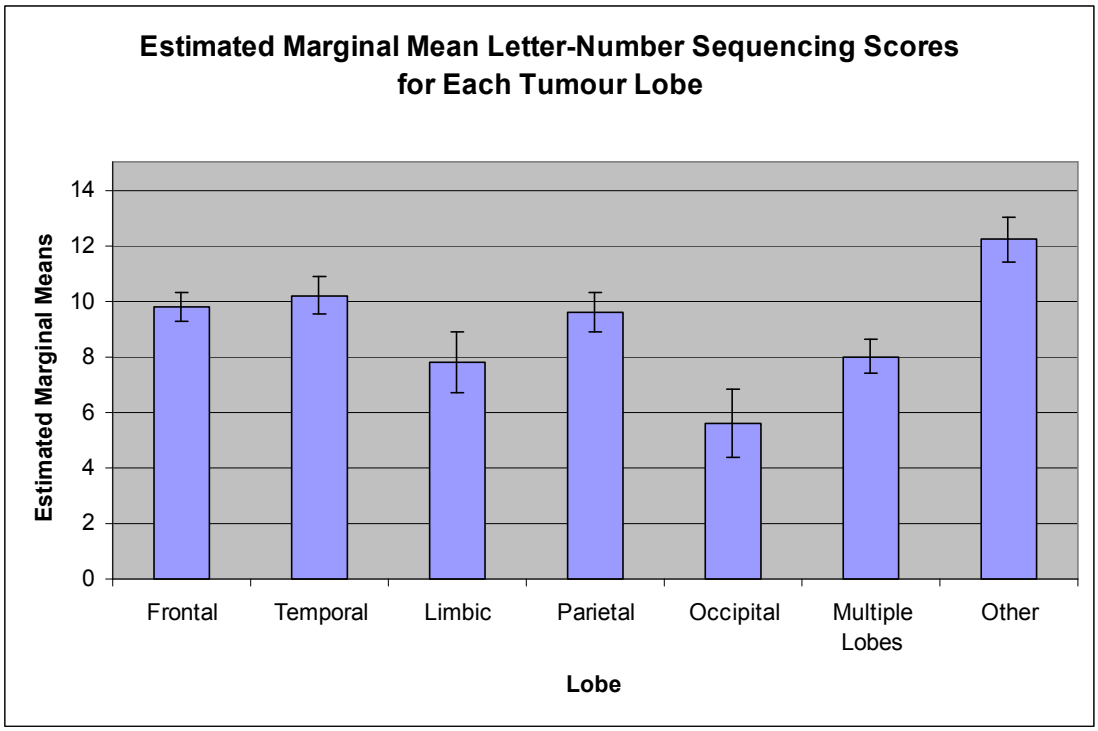


Figure 8.21. Baseline letter-number sequencing test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.8 Williams Delayed Recall Test (EFIT)

Thirty-four patients with frontal lobes tumours, 16 with temporal lobe tumours, 8 with limbic tumours, 17 with parietal tumours, 5 with occipital lobe tumours, 22 with tumours in multiple lobes and 13 with tumours located in 'other' brain areas completed this subtest from the Edinburgh Functional Impairment Tests.

The mean score for the frontal lobe group was 8.3 (SD 6.0) and for the temporal lobe group was 7.3 (SD 3.1). The limbic group scored a mean of 8.8 (SD 2.9), the parietal lobe group mean was 8.1 (SD 4.7), the occipital lobe group mean score was 14.6 (SD 6.3), the multiple lobes group had a mean score of 8.1 (SD 5.5) and the 'other' location group had a mean score of 6.5 (SD 3.6).

There was a significant main effect of the covariate age in the model, $F(1,99) = 13.246$, $p < 0.001$, partial $\eta^2 = 0.118$. Older participants were more likely to have higher (poorer) scores on this test. The covariate NART score also had a significant

main effect in the model, $F(1,99) = 5.336$, $p = 0.023$, partial $\eta^2 = 0.051$. Participants with lower (better) NART scores were significantly more likely to have lower (better) scores on this subtest. There was no significant effect of sex in the model, $F(1,99) = 2.532$, $p = 0.115$, partial $\eta^2 = 0.025$. The effect of tumour lobe was not significant in the model that included the effects of the covariates age and NART score, $F(6,99) = 1.740$, $p = 0.120$, partial $\eta^2 = 0.095$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.22, for each tumour lobe group in the Williams delayed recall test, were 8.4 (SE 0.8) for the frontal lobe tumour group; 7.3 (SE 1.2) for the temporal lobe group; 9.8 (SE 1.8) for the limbic group; 8.2 (SE 1.2) for the parietal lobe group; 14.0 (SE 2.2) for the occipital lobe group; 8.1 (SE 1.0) for the multiple lobe group and 6.3 (SE 1.3) for those patients with tumours located elsewhere in the brain ('other').

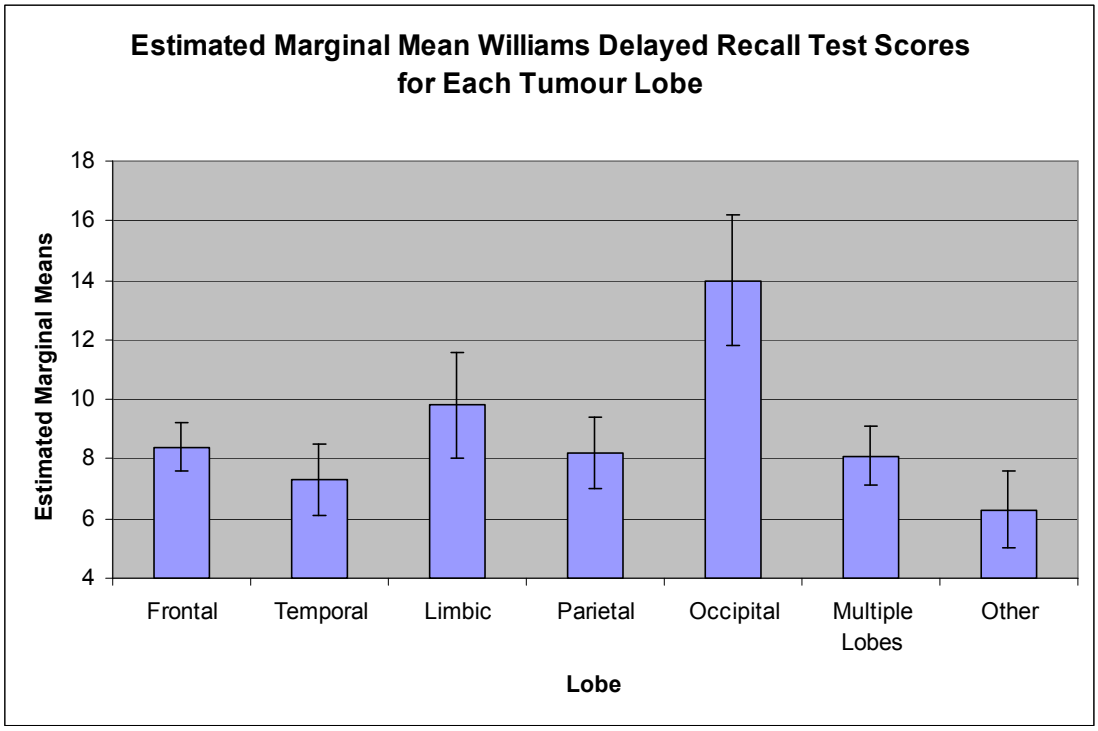


Figure 8.22. Baseline Williams Delayed Recall Test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.9 Nine Hole Peg Test (Right Hand, EFIT)

Thirty-four patients in the frontal lobe tumour group, 16 temporal lobe tumour patients, 8 patients with limbic tumours, 17 with parietal tumours, 5 patients with occipital lobe tumours, 22 with tumours in multiple lobes and 13 with tumours in 'other' locations in the brain all completed this subtest.

The frontal lobe group had a mean score of 13.7 (SD 2.5) on the test, the temporal lobe group had a mean score of 13.4 (SD 2.3) and the limbic group had a mean score of 14.5 (SD 2.8). The mean for the parietal lobe group was 18.5 (SD 8.7), for the occipital lobe group was 14.8 (SD 4.4) and for the multiple lobe group was 14.2 (SD 3.4). The 'other' location group scored a mean of 15.2 (SE 3.2) on the right hand Nine Hole Peg Test.

There was a significant main effect of age in the model, $F(1,99) = 14.891$, $p < 0.001$, partial $\eta^2 = 0.131$. Older patients were significantly more likely to take longer to complete the subtest than younger patients. There was also a significant main effect of the covariate NART score in the model, $F(1,99) = 16.020$, $p < 0.001$, partial $\eta^2 = 0.139$. Participants with lower (better) NART scores were more likely to have lower (better) scores on the subtest. There was no significant effect of sex in the model, $F(1,99) = 0.888$, $p = 0.348$, partial $\eta^2 = 0.009$. The effect of tumour lobe was significant in the model that included the effects of age and NART score, $F(6,99) = 3.268$, $p = 0.006$, partial $\eta^2 = 0.165$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.23, for each tumour lobe group on the right hand Nine Hole Peg Test were 13.4 (SE 0.7) for the frontal lobe group, 13.2 (SE 1.0) for the temporal lobe group, 15.5 (SE 1.5) for the limbic group, 18.4 (SE 1.0) for the parietal lobe group, 13.8 (SE 1.8) for the occipital lobe group, 14.0 (SE 0.8) for the multiple lobe group and 14.7 (SE 1.1) for the 'other' location group. Pairwise comparisons using LSD tests revealed that the parietal lobe group were significantly slower to complete this test than the frontal lobe group ($p < 0.001$); the temporal lobe group ($p < 0.001$); the occipital lobe group ($p = 0.028$); the multiple lobe group ($p = 0.001$) and the 'other'

location group (p = 0.013). There were no other significant differences between the tumour lobe groups.

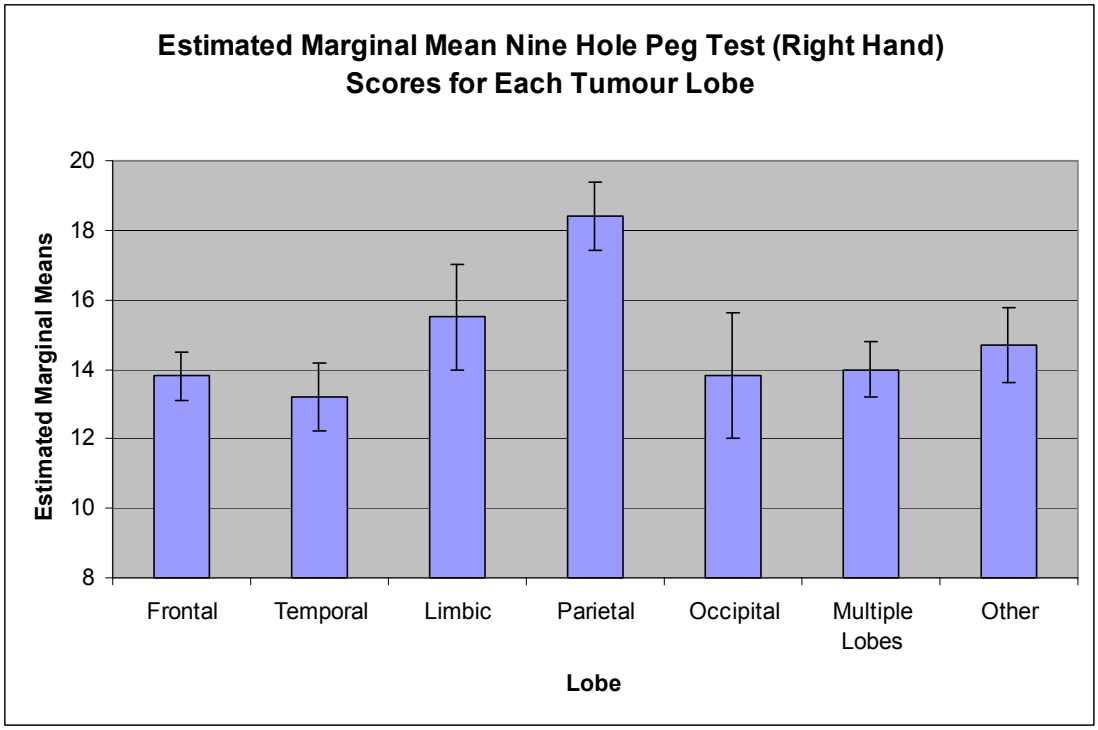


Figure 8.23. Baseline right hand nine hole peg test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.10 Nine Hole Peg Test (Left Hand, EFIT)

There were 33 patients in the frontal lobe group, 16 in the temporal lobe group, 8 with limbic tumours, 16 in the parietal lobe group, 5 in the occipital lobe group, 21 with tumours in multiple lobes and 13 with tumours in 'other' brain areas who completed this subtest.

The mean score for the frontal lobe tumour group was 15.3 (SD 4.3) and for the temporal lobe group was 14.2 (SD 2.3). The mean score for the limbic group was 14.6 (SD 2.5), for the parietal lobe group was 15.7 (SD 3.2) and for the occipital lobe group was 17.7 (SD 2.0). The multiple lobe group scored a mean of 15.5 (SD 3.8) and the 'other' location group score a mean of 16.3 (SD 3.4).

There was a significant main effect of the covariate age on test scores in the model, $F(1,96) = 18.648$, $p < 0.001$, partial $\eta^2 = 0.163$. Older participants were more likely to take longer to complete the subtest. There was no significant main effect of the covariate NART score, $F(1,96) = 1.889$, $p = 0.173$, partial $\eta^2 = 0.019$. The effect of sex was not significant in the model, $F(1,96) = 1.324$, $p = 0.253$, partial $\eta^2 = 0.014$. The effect of tumour lobe was not significant in the model that included the effects of age and NART score, $F(6,96) = 0.528$, $p = 0.786$, partial $\eta^2 = 0.032$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.24, for each tumour lobe group on the left hand nine hole peg test were as follows: 15.4 (SE 0.6) for the frontal lobe group; 14.2 (SE 0.8) for the temporal lobe group; 15.2 (SE 1.2) for the limbic group; 15.7 (SE 0.8) for the parietal group; 16.7 (SE 1.5) for the occipital group; 15.4 (SE 0.7) for the multiple lobe group and 16.0 (SE 0.9) for the 'other' location group.

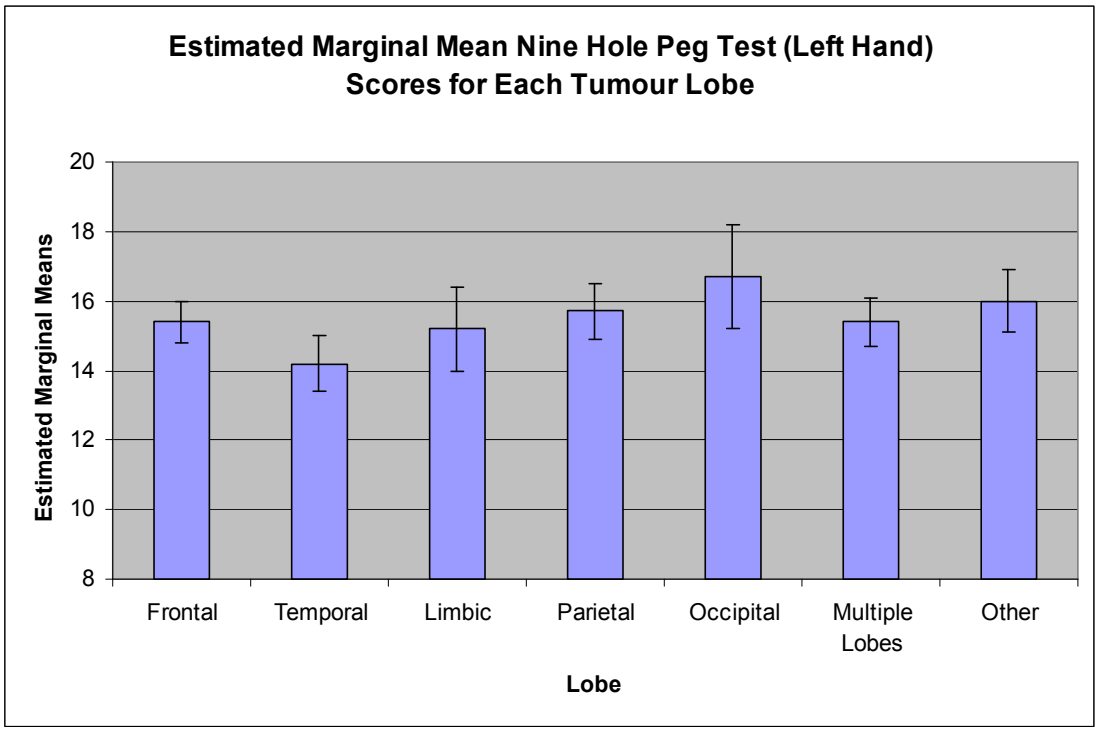


Figure 8.24. Baseline left hand nine hole peg test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.11 Timed Ten Metre Walk (EFIT)

Thirty-three patients in the frontal lobe tumour group and 15 in the temporal lobe group completed this test at baseline. Seven patients in the limbic group, 15 in the parietal lobe group, 3 in the occipital lobe group, 17 in the multiple lobe group and 12 in the 'other' location group also completed the subtest.

The frontal lobe group scored a mean of 6.9 (SD 1.4), the temporal lobe group had a mean score of 6.1 (SD 0.9) and the limbic group had a mean score of 7.6 (SD 2.9). The parietal lobe group had a mean score of 7.4 (SD 1.7), the occipital group mean was 6.2 (SD 0.5) and the multiple lobe group mean was 6.8 (SD 1.5). The 'other' tumour location group had a mean score of 7.9 (SD 2.8) on the timed ten metre walk subtest at baseline.

The covariate age had a significant main effect in the model, $F(1,86) = 14.032$, $p < 0.001$, partial $\eta^2 = 0.140$. Older participants were significantly more likely to take longer to complete the timed ten metre walk. The effect of the covariate NART did not reach the conventional level of statistical significance, $F(1,86) = 3.169$, $p = 0.079$, partial $\eta^2 = 0.036$. Sex had no significant effect in the model, $F(1,86) = 0.062$, $p = 0.805$, partial $\eta^2 = 0.001$. The effect of tumour lobe was not significant in the model that included the effects of the covariates, $F(6,86) = 1.879$, $p = 0.094$, partial $\eta^2 = 0.116$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.25, for each tumour lobe group on the timed ten metre walk subtest were as follows: 6.9 (SE 0.3) for the frontal lobe group; 6.1 (SE 0.4) for the temporal lobe group; 7.2 (SE 0.6) for the limbic group; 7.3 (SE 0.4) for the parietal lobe group; 5.7 (SE 0.9) for the occipital lobe group; 6.8 (SE 0.4) for the multiple lobes group and 7.8 (SE 0.4) for the 'other' location group.

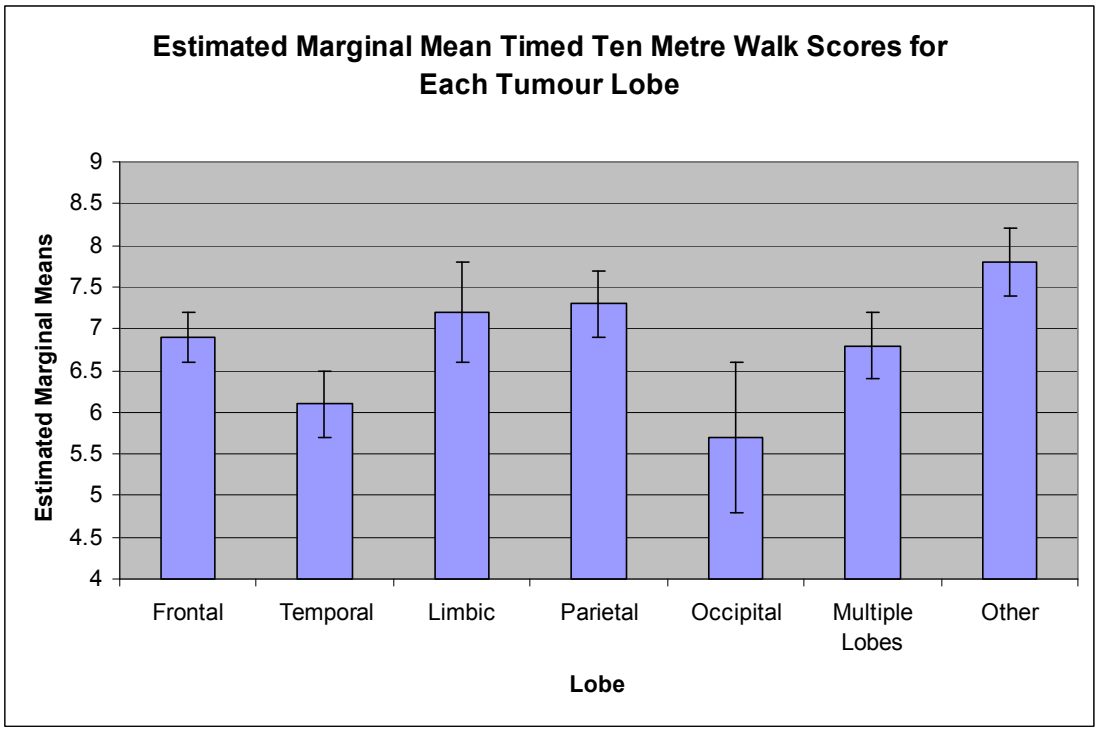


Figure 8.25. Baseline timed ten metre walk test scores for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score.
Error bars: Standard error means.

8.3.12 Hospital Anxiety and Depression Scale

There were 33 patients with frontal lobe tumours, 16 with temporal lobe tumours, 8 with limbic tumours, 18 with parietal lobe tumours, 5 with occipital lobe tumours, 22 with tumours in multiple lobes and 13 with tumours elsewhere in the brain ('other') who completed the Hospital Anxiety and Depression Scale at baseline.

8.3.12.1 Anxiety Scores

The frontal lobe group scored a mean anxiety score of 8.7 (SD 5.2) and the temporal lobe tumour group scored a mean of 8.2 (SD 5.4). The mean anxiety score for the limbic group was 8.0 (SD 4.4), for the parietal lobe group was 7.1 (SD 3.2), for the occipital lobe group was 9.6 (SD 3.8), for the multiple lobes group was 7.0 (SD 3.7) and for the 'other' location group was 8.6 (SD 3.5).

There was no significant main effect of the covariate age in the model, $F(1,99) = 0.020$, $p = 0.889$, partial $\eta^2 < 0.001$. The covariate NART score also had no significant main effect in the model, $F(1,99) = 0.186$, $p = 0.667$, partial $\eta^2 = 0.002$. Sex had no significant effect on anxiety scores in the model, $F(1,99) = 0.225$, $p = 0.636$, partial $\eta^2 = 0.002$. Tumour lobe had no significant main effect in the model that included the effects of the covariates, $F(6,99) = 0.807$, $p = 0.567$, partial $\eta^2 = 0.047$.

The estimated marginal mean scores, adjusted for age and NART score and shown in Figure 8.26, for each tumour lobe group on the anxiety measure of the Hospital Anxiety and Depression Scale were as follows: 8.7 (SE 0.8) for the frontal lobe group; 8.3 (SE 1.1) for the temporal lobe group; 7.4 (SE 1.6) for the limbic group; 6.9 (SE 1.1) for the parietal lobe group; 10.2 (SE 2.0) for the occipital lobe group; 7.0 (SE 0.9) for the multiple lobes group and 8.5 (SE 1.2) for the 'other' location group.

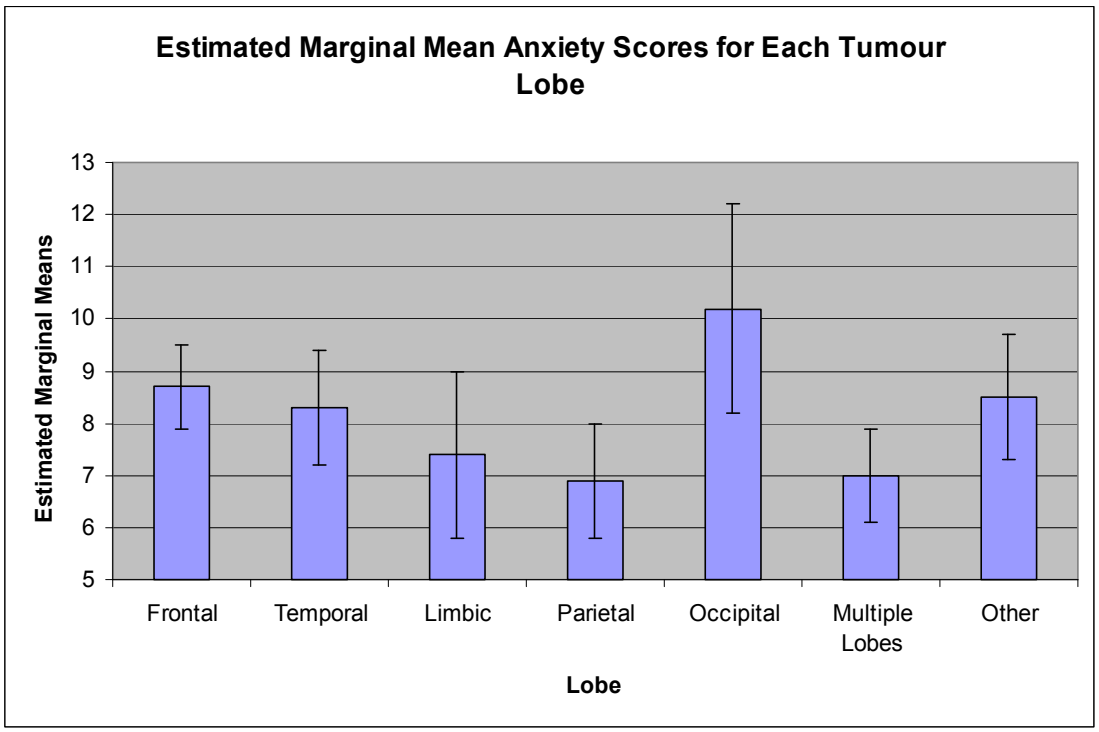


Figure 8.26. Baseline anxiety scores on the Hospital Anxiety and Depression Scale for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.3.12.2 Depression Scores

The mean depression score for the frontal lobe group was 5.6 (SD 4.7) and for the temporal lobe group was 4.3 (SD 3.9). The limbic group had a mean depression score of 3.8 (SD 2.6), the parietal group had a mean of 3.9 (SD 3.8), the occipital group had a mean of 4.2 (SD 4.6), the multiple lobes group had a mean of 3.5 (SD 3.3) and the 'other' location group had a mean score of 6.8 (SD 4.3).

There was no significant main effect of the covariate age in the model, $F(1,99) = 0.558$, $p = 0.457$, partial $\eta^2 = 0.006$. The covariate NART score also had no significant effect, $F(1,99) = 1.562$, $p = 0.214$, partial $\eta^2 = 0.016$. Sex had no significant effect on depression scores, $F(1,99) = 0.059$, $p = 0.808$, partial $\eta^2 = 0.001$. The effect of tumour lobe was not significant in the model that included the effects of the covariates, $F(6,99) = 1.352$, $p = 0.242$, partial $\eta^2 = 0.076$.

The estimated marginal mean depression scores, adjusted for age and NART score and shown in Figure 8.27, for each tumour lobe group were 5.6 (SE 0.7) for the frontal lobe group, 4.2 (SE 1.0) for the temporal lobe group, 4.1 (SE 1.4) for the limbic group, 3.7 (SE 1.0) for the parietal lobe group, 4.9 (SE 1.8) for the occipital lobe group, 3.5 (SE 0.8) for the multiple lobes group and 6.4 (SE 1.1) for the 'other' location group.

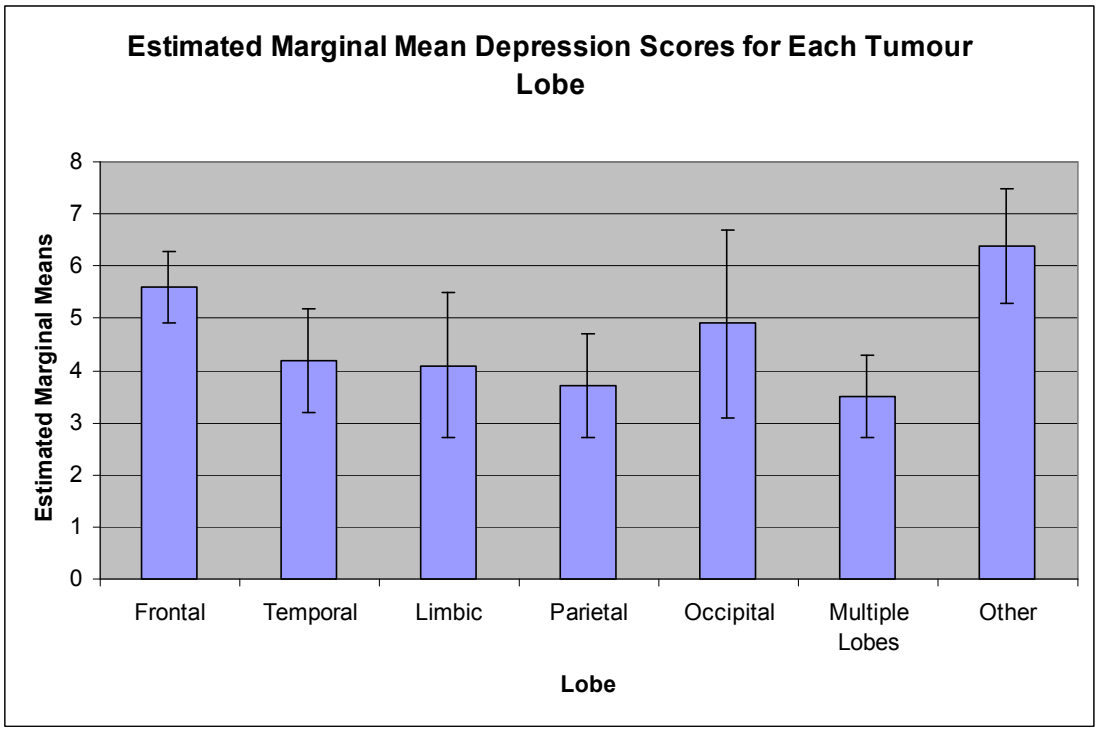


Figure 8.27. Baseline depression scores on the Hospital Anxiety and Depression Scale for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

8.3.12.3 Hospital Anxiety and Depression Scale – Total Score

The mean total score on the Hospital Anxiety and Depression Scale was 14.2 (SD 8.9) for the frontal lobe group, 12.5 (SD 8.2) for the temporal lobe group, 11.8 (SD 6.1) for the limbic group, 11.0 (SD 6.3) for the parietal lobe group, 13.8 (SD 8.3) for the occipital lobe group, 10.5 (SD 6.2) for the multiple lobe group and 15.4 (SD 6.7) for the group with tumours located elsewhere in the brain ('other').

There was no significant main effect of the covariate age in the model, $F(1,99) = 0.229$, $p = 0.634$, partial $\eta^2 = 0.002$. The covariate NART score also had no significant effect in the model, $F(1,99) = 0.841$, $p = 0.361$, partial $\eta^2 = 0.008$. Sex had no significant effect on total scores in the model, $F(1,99) = 0.023$, $p = 0.879$, partial $\eta^2 < 0.001$. The effect of tumour lobe was not significant in the model that included the effects of age and NART score, $F(6,99) = 1.163$, $p = 0.333$, partial $\eta^2 = 0.066$.

The estimated marginal mean total Hospital Anxiety and Depression Scale scores, adjusted for age and NART score and shown in Figure 8.28, for each tumour lobe group were 14.4 (SE 1.3) for the frontal lobe group, 12.6 (SE 1.9) for the temporal lobe group, 11.5 (SE 2.7) for the limbic group, 10.5 (SE 1.8) for the parietal lobe group, 15.1 (SE 3.4) for the occipital lobe group, 10.5 (SE1.6) for the multiple lobe group and 14.9 (SE 2.1) for the 'other' location group.

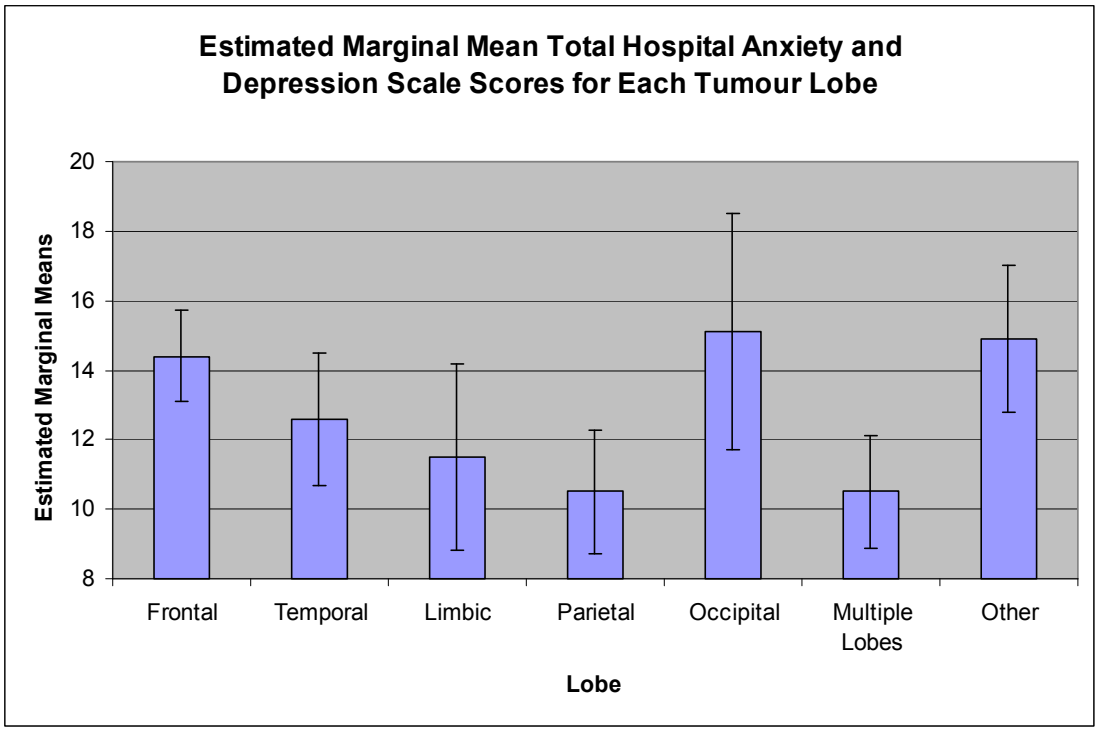


Figure 8.28. Baseline total scores on the Hospital Anxiety and Depression Scale for each tumour lobe group. Bars are estimated marginal mean scores adjusted for age and NART score. Error bars: Standard error means.

Table 8.2. Comparison of tumour lobe groups on cognitive and mood scales at baseline.

Test	Effect of tumour lobe			Overview of significant pairwise comparisons
	F	Sig. (p)	n^{2}*	
Inspection Time (All Data)	3.372	0.002	0.189	Occipital lobe group performed less well than all other lobe groups.
Inspection Time (Valid Data)	1.268	0.281	0.086	N/A
Rey Auditory Verbal Learning Test (Total)	2.268	0.046	0.161	'Other' lobe group performed better than temporal, limbic, occipital and multiple lobe groups.
Trail Making Test Part B (secs)	0.846	0.538	0.060	N/A
Verbal Fluency (Total)	1.124	0.357	0.083	N/A
Digit Symbol Coding (Total)	2.180	0.052	0.121	Frontal group performed better than parietal and temporal. 'Other' group performed better than parietal. Temporal group performed better than multiple lobes group.

Letter-Number Sequencing (Total)	5.166	0.001	0.298	'Other' group performed better than frontal, limbic, parietal, occipital and multiple lobe groups. Occipital group performed less well than temporal, parietal, frontal lobe groups. Temporal group performed better than multiple lobes group.
EFIT Williams Delayed Recall Test (total)	1.740	0.120	0.095	N/A
EFIT Nine Hole Peg Test (Right Hand, secs)	3.268	0.006	0.165	Parietal group performed less well than frontal, temporal, occipital, multiple and 'other' lobe groups.
EFIT Nine Hole Peg Test (Left Hand, secs)	0.786	0.786	0.032	N/A
EFIT Timed Ten Metre Walk (secs)	1.879	0.094	0.116	N/A
Hospital Anxiety and Depression Scale –Anxiety Score	0.807	0.567	0.047	N/A

Hospital Anxiety and Depression Scale – Depression Score	1.352	0.242	0.076	N/A
Hospital Anxiety and Depression Scale – Total Score	1.163	0.333	0.066	N/A

n^2 = the proportion of variance accounted for by tumour lobe.

8.4 Hemispheric Lateralisation

Of the total cohort of brain tumour patients (n = 118), 53 had a tumour located within the left hemisphere and 53 had a tumour located within the right hemisphere. The remaining twelve patients had either bi-hemispheric tumours, tumours located on the pituitary gland or tumours located between the ventricles, and were not included in this analysis.

8.4.1 Inspection Time Scores: All Inspection Time Data

The mean baseline score for the patients with tumours located in the left hemisphere was 118.6 (SD 14.9). Patients with tumours in the right hemisphere had a mean baseline inspection time score of 109.2 (SD 23.7).

The covariate age had a significant effect in this model, $F(1,100) = 43.254$, $p < 0.001$, partial $\eta^2 = 0.302$. Older patients were more likely to have poorer inspection time scores. The covariate NART score also had a significant effect, $F(1,100) = 11.525$, $p = 0.001$, partial $\eta^2 = 0.103$. Those participants with lower (better) NART scores were significantly more likely to have higher (better) inspection time scores. The effect of sex was not significant, $F(1,100) = 3.332$, $p = 0.071$, partial $\eta^2 = 0.032$. Hemispheric location of tumour had a significant main effect in the model that included the effects of the covariates age and NART score, $F(1,100) = 10.987$, $p = 0.001$, partial $\eta^2 = 0.099$.

The estimated marginal mean scores on this test were 119.5 (SE 2.2) for the left hemisphere tumour group and 109.0 (SE 2.2) for the right hemisphere tumour group. The significant effect of hemisphere in the model therefore suggests that patients with a left-sided brain tumour had significantly better inspection time performance than the group of patients who had a tumour located in the right hemisphere.

8.4.2 Inspection Time Scores: Valid Inspection Time Data

When only data from patients with 'valid' inspection time scores was included in the analyses, this resulted in the left hemisphere tumour group including data from 50 patients, with 40 patients with tumours in the right hemisphere.

When only valid scores were included, the left hemisphere tumour group scored a mean of 120.2 (SD 13.4) on the inspection time measure at baseline, and the right hemisphere group scored a mean of 119.7 (SD 14.7).

There was again a significant effect of age in the model, $F(1,84) = 27.763$, $p < 0.001$, partial $\eta^2 = 0.248$. Older participants had poorer scores on the inspection time measure. The effect of the covariate NART score approached the conventional level of statistical significance on this occasion, $F(1,84) = 3.725$, $p = 0.057$, partial $\eta^2 = 0.042$. Sex also had no significant effect in the model, $F(1,84) = 3.137$, $p = 0.080$, partial $\eta^2 = 0.036$. In the model that controlled for the effects of age and NART score, hemisphere had no significant effect when only 'valid' inspection time data was included, $F(1,84) = 1.716$, $p = 0.194$, partial $\eta^2 = 0.020$.

The estimated marginal mean score, adjusted for age and NART score, for the left hemisphere tumour group was 122.0 (SE 1.7) and for the right hemisphere group was 118.6 (SE 1.9). When the invalid inspection time scores were removed from the model, there were no significant differences between the baseline inspection time scores of the left and the right hemisphere tumour groups.

8.4.3 Rey Auditory Verbal Learning Test

Thirty-eight patients with tumours in the left hemisphere and 43 with tumours in the right hemisphere completed this measure at baseline.

The mean score for the left hemisphere group was 57.2 (SD 16.1) and for the right hemisphere group was 65.7 (SD 18.3).

There was a significant main effect of the covariate age in the model, $F(1,75) = 25.180$, $p < 0.001$, partial $\eta^2 = 0.251$. Older patients were significantly more likely to have lower scores on this measure. The effect of the covariate NART score was also significant, $F(1,75) = 11.935$, $p = 0.001$, partial $\eta^2 = 0.137$. Lower (better) NART scores were significantly associated with higher (better) scores on the Rey Auditory Verbal Learning Test. Sex also had a significant main effect in the model, $F(1,75) = 6.824$, $p = 0.011$, partial $\eta^2 = 0.083$. Female participants tended to have higher scores on this test than their male counterparts. The effect of tumour hemisphere was significant in the model that included the effects of age and NART score, $F(1,75) = 6.778$, $p = 0.011$, partial $\eta^2 = 0.083$.

The estimated marginal mean scores, adjusted for age and NART score, on the Rey Auditory Verbal Learning Test, were 57.4 (SE 2.4) for the left hemisphere tumour group and 65.8 (SE 2.2) for the right hemisphere tumour group. Therefore, the group of patients with tumours located within the left hemisphere performed significantly less well than the group of patients with right hemisphere tumours.

8.4.4 Trail Making Test Part B

Forty-six patients with tumours in the left hemisphere and 45 with right hemisphere tumours completed the trail making test part B at baseline.

The mean score for the left hemisphere group was 96.3 (SD 33.2) and for the right hemisphere group was 93.0 (SD 28.5).

There was a significant main effect of the covariate age in the model, $F(1,85) = 19.598$, $p < 0.001$, partial $\eta^2 = 0.187$. Older participants were more likely to take longer to complete the trial making test part B than younger participants. The covariate NART score also had a significant main effect in the model, $F(1,85) = 22.754$, $p < 0.001$, partial $\eta^2 = 0.211$. Poorer performance on the NART was significantly associated with slower performance on the trail making test part B. The

effect of sex was not significant in the model, $F(1,85) = 0.067$, $p = 0.797$, partial $\eta^2 = 0.001$. The hemispheric location of the tumour had no significant effect in the model that included the effects of age and NART score, $F(1,85) = 0.003$, $p = 0.960$, partial $\eta^2 < 0.001$.

The estimated marginal mean scores, adjusted for age and NART score, for the trail making test part B were 93.8 (SE 3.9) for the left hemisphere group and 94.1 (SE 3.9) for the right hemisphere group. There was no significant difference between the performance of those patients with left hemisphere tumours and those patients with right hemisphere tumours on the trail making test part B.

8.4.5 Verbal Fluency

Forty patients with left hemisphere tumours and 46 with right hemisphere tumours completed the verbal fluency test at baseline.

The mean score for the left hemisphere tumour group was 28.3 (SD 9.8) and for the right hemisphere tumour group was 28.7 (SD 14.1).

There was a significant main effect of the covariate age in the model, $F(1,80) = 6.070$, $p = 0.016$, partial $\eta^2 = 0.071$. Older patients tended to produce fewer words on the verbal fluency task than younger participants. There was a significant main effect of the covariate NART score in the model, $F(1,80) = 24.142$, $p < 0.001$, partial $\eta^2 = 0.232$. Higher (poorer) NART scores were significantly associated with lower (poorer) verbal fluency scores. The effect of sex did not reach the conventional level of statistical significance, $F(1,80) = 3.315$, $p = 0.072$, partial $\eta^2 = 0.040$. The effect of hemispheric location of tumour was not significant in the model that included the effects of age and NART score, $F(1,80) = 0.079$, $p = 0.780$, partial $\eta^2 = 0.001$.

The estimated marginal mean scores, adjusted for age and NART score, for each hemisphere group on the verbal fluency measure were 28.6 (SD 1.7) for the left hemisphere tumour group and 29.2 (SD 1.5) for the right hemisphere group. There

was no difference between the performance of the two groups on the verbal fluency measure.

8.4.6 Digit Symbol Coding

There were 54 patients with a right hemisphere tumour and 52 with a left hemisphere tumour who completed the digit symbol coding task at baseline.
The mean score for the left hemisphere group was 57.4 (SD 21.0) and for the right hemisphere group was 54.4 (SD 22.9).

There was a significant main effect of the covariate age in the model, $F(1,100) = 51.387$, $p < 0.001$, partial $\eta^2 = 0.339$. Older participants were significantly more likely to have lower (poorer) scores on this test. The covariate NART also had a significant main effect in the model, $F(1,100) = 34.364$, $p < 0.001$, partial $\eta^2 = 0.256$. Those participants with higher (poorer) NART sores were significantly more likely to perform less well on the digit symbol coding measure. There was no significant effect of sex in the model, $F(1,100) = 0.662$, $p = 0.418$, partial $\eta^2 = 0.007$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates age and NART score, $F(1,100) = 1.177$, $p = 0.281$, partial $\eta^2 = 0.012$.

The estimated marginal mean scores, adjusted for age and NART score, for each hemisphere group on the digit symbol coding test at baseline were 57.7 (SE 2.3) for the left hemisphere tumour group and 54.1 (SE 2.4) for the right hemisphere tumour group. There was no significant effect of the hemispheric location of the tumour on digit symbol coding scores.

8.4.7 Letter-Number Sequencing

Forty-one patients with left hemisphere tumours and 43 patients with right hemisphere tumours completed the letter-number sequencing test at baseline.

The mean score for the left hemisphere group was 9.3 (SD 3.4) and for the right hemisphere group was 9.5 (SD 3.4).

There was a significant main effect of both covariates age, $F(1,78) = 29.787$, $p < 0.001$, partial $\eta^2 = 0.276$; and NART score, $F(1,78) = 26.102$, $p < 0.001$, partial $\eta^2 = 0.251$, in the model. Older patients and patients with poorer scores on the NART were more likely to have poorer scores on the letter-number sequencing test. There was no significant effect of sex on test scores in the model, $F(1,78) = 0.375$, $p = 0.542$, partial $\eta^2 = 0.005$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates age and NART score, $F(1,78) = 0.001$, $p = 0.971$, partial $\eta^2 < 0.001$.

The estimated marginal mean scores, adjusted for age and NART score, on the letter-number sequencing test at baseline for each tumour hemisphere group were 9.4 (SE 0.4) for the left hemisphere tumour group and 9.5 (SE 0.4) for the right hemisphere tumour group. The hemisphere in which a patient's tumour was located had no effect on letter-number sequencing scores at baseline.

8.4.8 Williams Delayed Recall Test (EFIT)

Fifty-four patients with left hemisphere tumours and 55 patients with right hemisphere tumours completed the Williams delayed recall test at baseline.

The mean score for the left hemisphere group was 8.0 (SD 4.4) and the mean score for the right hemisphere group was 8.2 (SD 5.9).

There was a significant main effect of both covariates age, $F(1,103) = 15.131$, $p < 0.001$, partial $\eta^2 = 0.128$; and NART score, $F(1,103) = 4.251$, $p = 0.042$, partial $\eta^2 = 0.040$, in the model. Older patients and patients with higher (poorer) NART scores were significantly more likely to have higher (poorer) scores on the Williams delayed recall test. The effect of sex was not significant, $F(1,103) = 0.740$, $p = 0.392$,

partial $\eta^2 = 0.007$. Tumour hemisphere had no significant effect in the model that included the effects of both covariates, $F(1,103) = 0.001$, $p = 0.975$, partial $\eta^2 < 0.001$.

The estimated marginal mean scores, adjusted for age and NART score, on the Williams delayed recall test were 8.1 (SE 0.7) for the left hemisphere group and 8.1 (SE 0.7) for the right hemisphere group. The hemispheric location of the tumour had no effect on how well the patient performed on the Williams delayed recall test.

8.4.9 Nine Hole Peg Test (Right Hand, EFIT)

There were 54 patients with left hemisphere tumours and 55 patients with right hemisphere tumours who completed this test at baseline.
The left hemisphere tumour group had a mean score of 15.2 (SD 5.7) and the right hemisphere tumour group had a mean score of 14.1 (SD 3.1) on the right-hand nine hole peg test.

There was a significant main effect of the covariate age in the model, $F(1,103) = 16.277$, $p < 0.001$, partial $\eta^2 = 0.136$. Older participants were more likely to perform more slowly on this test. The covariate NART score also had a significant main effect in the model, $F(1,103) = 15.229$, $p < 0.001$, partial $\eta^2 = 0.129$. Higher (poorer) NART scores were significantly associated with slower performance on this test. There was no significant effect of sex in the model, $F(1,103) = 0.996$, $p = 0.321$, partial $\eta^2 = 0.010$. There was no significant effect of hemispheric location of tumour in the model that included the effects of age and NART score, $F(1,103) = 2.126$, $p = 0.148$, partial $\eta^2 = 0.020$.

The estimated marginal mean scores, adjusted for age and NART score, on the right hand nine hole peg test were 15.3 (SE 0.6) for the left hemisphere tumour group and 14.1 (SD 0.6) for the right hemisphere tumour group. Therefore, there was no significant difference between those patients with left hemisphere tumours and those with right hemisphere tumours in terms of right hand nine hole peg test performance.

8.4.10 Nine Hole Peg Test (Left Hand, EFIT)

Fifty-four patients with left hemisphere tumours and 52 with right hemisphere tumours completed this test at baseline.

The mean score for the left hemisphere tumour group was 14.6 (SD 2.5) and for the right hemisphere tumour group was 16.2 (SD 4.2).

The covariate age had a significant main effect in the model, $F(1,100) = 20.376$, $p < 0.001$, partial $\eta^2 = 0.169$. Older patients were more likely to take longer to complete this test than younger patients. The effect of the covariate NART score approached the conventional level of statistical significance, $F(1,100) = 3.600$, $p = 0.061$, partial $\eta^2 = 0.035$. Sex had no significant effect in the model, $F(1,100) = 1.976$, $p = 0.163$, partial $\eta^2 = 0.019$. The effect of tumour hemisphere was significant in the model that included the effects of age and NART score, $F(1,100) = 5.653$, $p = 0.019$, partial $\eta^2 = 0.054$.

The estimated marginal mean scores, adjusted for age and NART score, on the left hand nine hole peg test were 14.7 (SE 0.4) for the left hemisphere tumour group and 16.2 (SE 0.4) for the right hemisphere group. Patients with tumours located in the right hemisphere were significantly slower to complete the left hand nine hole peg test than those patients with left hemisphere tumours.

8.4.11 Timed Ten Metre Walk (EFIT)

Forty-nine patients with left hemisphere tumours and 47 with right hemisphere tumours completed the timed ten metre walk at baseline.

The mean score for the left hemisphere group was 6.8 (SD 1.9) and for the right hemisphere group was 7.0 (SD 1.4).

There was a significant main effect of age in the model, $F(1,90) = 19.775$, $p < 0.001$, partial $\eta^2 = 0.180$. Older patients took significantly longer to complete the timed ten

metre walk than younger patients. The effect of the covariate NART score almost reached the conventional level of statistical significance, $F(1,90) = 3.858$, $p = 0.053$, partial $\eta^2 = 0.041$. The effect of sex was not significant in the model, $F(1,90) = 2.464$, $p = 0.120$, partial $\eta^2 = 0.027$. The effect of tumour hemisphere was not significant in the model that included the effects of age and NART score, $F(1,90) = 0.616$, $p = 0.435$, partial $\eta^2 = 0.007$.

The estimated marginal mean scores, adjusted for age and NART score, on the timed ten metre walk were 6.8 (SE 0.2) for the left hemisphere tumour group and 7.0 (SE 0.2) for the right hemisphere tumour group. The hemispheric location of tumour had no significant effect on timed ten metre walk test performance.

8.4.12 Hospital Anxiety and Depression Scale

Fifty-five patients with tumours located in the left hemisphere and 54 patients with tumours located in the right hemisphere completed the Hospital Anxiety and Depression Scale at baseline.

8.4.12.1 Anxiety Scores

The mean anxiety score for the left hemisphere tumour group was 8.0 (SD 4.4) and for the right hemisphere tumour group was 7.9 (SD 4.4).

There was no significant effect of either covariates age, $F(1,103) = 0.018$, $p = 0.894$, partial $\eta^2 < 0.001$; or NART score, $F(1,103) = 0.057$, $p = 0.812$, partial $\eta^2 = 0.001$, in the model. There was no significant effect of sex, $F(1,103) = 1.224$, $p = 0.271$, partial $\eta^2 = 0.012$. The effect of tumour hemisphere was not significant in the model that included the effects of age and NART score, $F(1,103) = 0.010$, $p = 0.919$, partial $\eta^2 < 0.001$.

The estimated marginal mean anxiety scores, adjusted for age and NART score, were 7.9 (SE 0.6) for the left hemisphere group and 7.9 (SE 0.6) for the right hemisphere

group. Hemispheric location of tumour had no significant effect on anxiety levels, as measured by the Hospital Anxiety and Depression Scale.

8.4.12.2 Depression Scores

The mean depression score for the left hemisphere tumour group was 4.8 (SD 4.4) and for the right hemisphere tumour group was 4.4 (SD 3.7).

There was no significant main effect of either of the covariates age, $F(1,103) = 0.135$, $p = 0.714$, partial $\eta^2 = 0.001$; or NART score, $F(1,103) = 0.783$, $p = 0.378$, partial $\eta^2 = 0.008$. The effect of sex approached the conventional level of statistical significance, $F(1,103) = 3.686$, $p = 0.0.58$, partial $\eta^2 = 0.035$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates, $F(1,103) = 0.016$, $p = 0.900$, partial $\eta^2 < 0.001$.

The estimated marginal mean depression scores, adjusted for age and NART score, were 4.5 (SE 0.6) for the left hemisphere tumour group and 4.4 (SE 0.6) for the right hemisphere tumour group. Hemispheric location of tumour had no significant effect on depression levels, as measured by the Hospital Anxiety and Depression Scale.

8.4.12.3 Hospital Anxiety and Depression Scale - Total Scores

The left hemisphere group had a mean total score of the Hospital Anxiety and Depression Scale of 12.8 (SD 8.0) and the right hemisphere tumour group had a mean of 12.3 (SD 7.1).

There was no significant main effect of either of the covariates age, $F(1,103) = 0.076$, $p = 0.784$, partial $\eta^2 = 0.001$; or NART score, $F(1,103) = 0.377$, $p = 0.541$, partial $\eta^2 = 0.004$, on total scores in the model. The effect of sex was not significant, $F(1,103) = 2.821$, $p = 0.096$, partial $\eta^2 = 0.027$. Tumour hemisphere had no significant effect on total scores in the model that included the effects of age and NART score, $F(1,103) < 0.001$, $p = 0.994$, partial $\eta^2 < 0.001$.

The estimated marginal mean total Hospital Anxiety and Depression Scale scores, adjusted for age and NART score, were 12.4 (SE 1.0) for the left hemisphere tumour group and 12.3 (SD 1.0) for the right hemisphere tumour group. Therefore, the hemispheric location of the tumour has no effect on the levels of distress as measured by the Hospital Anxiety and Depression Scale.

The results detailed in this chapter are discussed in Chapter 9.6.

Table 8.3. Comparison of the left and right hemisphere groups on cognitive and mood scales at baseline.

Test	Effect of tumour hemisphere			Estimated marginal mean score (adjusted for age and NART score)	
	F	Sig. (p)	n^{2*}	Left Hemisphere	Right Hemisphere
Inspection Time (All Data)	10.987	0.001	0.099	119.5 (2.2)	109.0 (2.2)
Inspection Time (Valid Data)	1.716	0.194	0.020	122.0 (1.7)	118.6 (1.9)
Rey Auditory Verbal Learning Test (Total)	6.778	0.011	0.083	57.4 (2.4)	65.8 (2.2)
Trail Making Test Part B (secs)	0.003	0.960	<0.001	93.8 (3.9)	94.1 (3.9)
Verbal Fluency (Total)	0.079	0.780	0.001	28.6 (1.7)	29.2 (1.5)
Digit Symbol Coding (Total)	1.177	0.281	0.012	57.7 (2.3)	54.1 (2.4)
Letter-Number Sequencing (Total)	0.001	0.971	<0.001	9.4 (0.4)	9.5 (0.4)
EFIT Williams Delayed Recall Test (total)	0.001	0.975	<0.001	8.1 (0.7)	8.1 (0.7)
EFIT Nine Hole Peg Test (Right Hand, secs)	2.216	0.148	0.020	15.3 (0.6)	14.1 (0.6)
EFIT Nine Hole Peg Test (Left Hand, secs)	5.653	0.019	0.054	14.7 (0.4)	16.2 (0.4)
EFIT Timed Ten Metre Walk (secs)	0.616	0.435	0.007	6.8 (0.2)	7.0 (0.2)
Hospital Anxiety and Depression Scale –Anxiety Score	0.010	0.919	<0.001	7.9 (0.6)	7.9 (0.6)

Hospital Anxiety and Depression Scale – Depression Score	0.016	0.900	<0.001	4.5 (0.6)	4.4 (0.6)
Hospital Anxiety and Depression Scale – Total Score	<0.001	0.994	<0.001	12.4 (1.0)	12.3 (1.0)

* n^2 = the proportion of variance accounted for by the tumour hemisphere

9 Tumour Type, Location and Lateralisation: Post-Operative Function

9.1 Overview of analysis procedure

To investigate the effect of surgery on inspection time, digit-symbol coding and the Edinburgh Functional Impairment Tests dependent on tumour histology, location (lobe) and hemispheric laterality, general linear modelling (analysis of covariance) was used. The post-operative (session 2) test score was entered as the dependent variable and the relevant tumour-related variable (type, lobe or hemisphere) and sex were specified as fixed effects (between-groups factors) in each model. Age and National Adult Reading Test (NART) score were covariates in each model for the reasons highlighted in chapter 5.2.1. The corresponding baseline test score was also included as a covariate in each model as this allows us to compare any differences between baseline and session 2 performance on each test for each histological group, tumour lobe group and hemispheric location by controlling for the effects of baseline test score in addition to the effects of age and NART score.

In each of the following sections, mean baseline and session 2 scores for each histological, lobe and hemisphere group are presented for each test with standard deviations (SD) in brackets. The results of the general linear model are then presented. Estimated marginal mean scores and pairwise comparisons are then given for each group to highlight any differences between the groups, adjusted for other variables in the models when a main effect of the tumour-related variable is found. Least Significant Difference (LSD) tests were used to conduct pairwise post-hoc comparisons. Post-hoc analyses comparing the low-grade and high-grade tumour groups are carried out for each measure, since it was expected that there would be a significant difference between these two histological groups (see Chapter 1.3 and 1.5).

9.2 Histological Tumour Type

Of the total cohort of brain tumour patients who completed both baseline and session 2 (post-operative) testing, 16 patients were found to have a low-grade glioma, 26 had a high-grade glioma, 9 had a metastasis, 10 had a meningioma and 3 had a diagnosis that did not fall into the aforementioned categories ('other').

9.2.1 Inspection Time Scores: All Inspection Time Data

Two patients with a high-grade glioma and 2 with a meningioma were tested at session 2 but did not complete the inspection time task. Therefore, data from 16 low-grade glioma patients, 24 high-grade glioma patients, 9 metastasis, 8 meningioma and 3 patients with 'other' tumours were included in the following analysis.

The mean baseline inspection time score for the low-grade glioma group was 127.3 (SD 10.8) and at session 2 was 118.2 (SD 23.4). The high-grade glioma group had a mean score at baseline of 113.3 (SD 22.3) and 113.3 (SD 22.5) at session 2. The baseline mean score for the metastasis group was 120.3 (SD 17.0) and at session 2 was 111.0 (SD 23.0). The meningioma group had a mean baseline score of 106.6 (SD 30.2) and a mean of 103.6 (SD 22.5) at session 2. The 'other' tumour group mean at baseline was 122.0 (SD 12.5) and at session 2 was 132.3 (SD 5.8). The mean (standard error) scores at baseline and session 2 for each tumour group are shown in Figure 9.1.

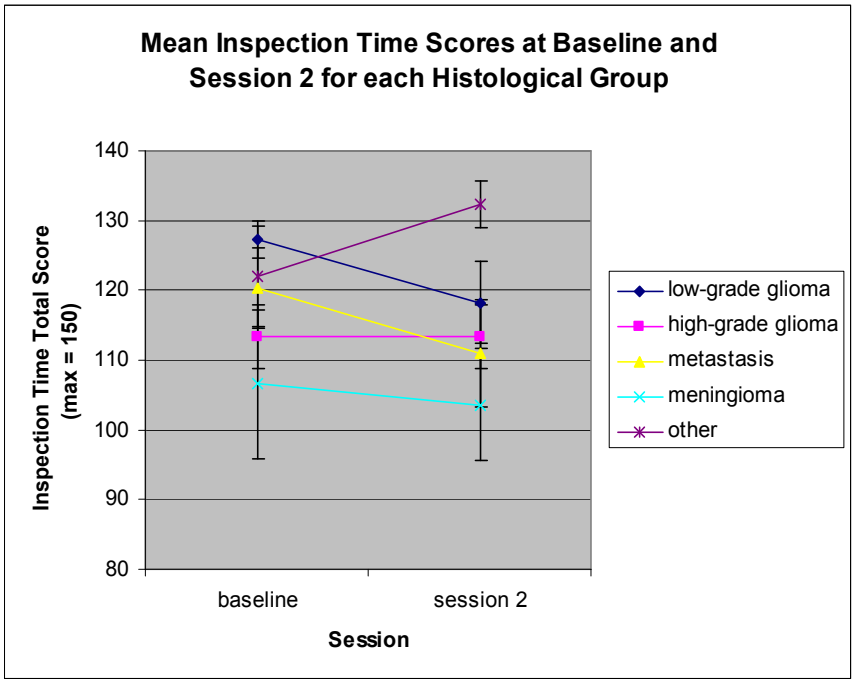


Figure 9.1. Baseline and session 2 inspection time scores for each histological group. Points are raw mean scores. Error bars: standard error means.

General linear modelling with post-operative (session 2) inspection time score entered as the dependent variable revealed a significant main effect of the baseline inspection time score, $F(1,48) = 28.663$, $p < 0.001$, partial $\eta^2 = 0.374$. In the whole sample (i.e. all histology groups combined), baseline inspection time score was positively correlated with session 2 inspection time score, $r(n = 60) = 0.709$, $p < 0.001$. Similar positive correlations were obtained for the high-grade glioma group, $r(n = 24) = 0.836$, $p < 0.001$; the metastasis group, $r(n = 9) = 0.849$, $p = 0.004$ and the meningioma group, $r(n = 8) = 0.819$, $p = 0.013$. The two scores were not significantly correlated in the low-grade glioma group, $r(n = 16) = 0.367$, $p = 0.162$ or in the 'other' tumour group, $r(n = 3) = 0.829$, $p = 0.377$. The effect of the covariate age was not significant in the model, $F(1,48) = 2.201$, $p = 0.144$, partial $\eta^2 = 0.044$. The effect of the covariate NART score was also not significant, $F(1,48) = 1.088$, $p = 0.302$, partial $\eta^2 = 0.022$. Sex had no significant main effect in the model, $F(1,48) = 0.207$, $p = 0.651$, partial $\eta^2 = 0.004$. The effect of tumour type was not significant in the model that included the effects of the covariates age, NART score and baseline inspection time score, $F(4,48) = 1.100$, $p =$

0.368, partial $\eta^2 = 0.084$. There was no significant interaction between tumour type and sex, $F(3,48) = 0.321$, $p = 0.810$, partial $\eta^2 = 0.020$.

The estimated marginal mean session 2 inspection time score, adjusted for age, NART and baseline inspection time score—derived from the above-described general linear model—was 109.6 (SE 4.3) for the low-grade glioma group; 117.2 (SE 3.5) for the high-grade glioma group; 111.3 (SE 5.8) for the metastasis group; 116.4 (SE 9.1) for the meningioma group and 126.7 (SE 9.4) for the 'other' tumour group. Pairwise comparisons using Least Significant Difference tests revealed no significant difference between the low-grade and high-grade glioma patient groups.

9.2.2 Inspection Time Scores: Valid Inspection Time Data

When only data from patients with 'valid' inspection time scores at baseline testing were included in the analysis, there were 16 low-grade glioma patients, 21 high-grade glioma patients, 8 metastasis patients, 5 meningioma patients and 2 patients with 'other' tumours who also had session 2 inspection time scores.

The mean baseline score for the low-grade glioma group was 127.3 (SD 10.8) and at session 2 was 118.2 (SD 23.4). The high-grade glioma group had a mean score of 119.2 (SD 15.1) at baseline and 116.8 (SD 17.7) at session 2. The mean baseline score for the metastasis group was 125.0 (SD 10.2) and at session 2 was 114.9 (SD 21.2). The meningioma group had a baseline mean score of 127.8 (SD 8.7) and 115.6 (SD 20.1) at session 2. The 'other' tumour type group had a mean baseline score of 128.5 (SD 7.8) and 134.0 (SD 7.1) at session 2. The mean (standard error) scores at baseline and session 2 for each tumour type group are shown in Figure 9.2.

Figure 9.2. Baseline and session 2 inspection time scores, with only valid baseline scores included, for each tumour type group. Points are raw mean scores. Error bars: standard error means.

There was a significant main effect of the inspection time score at baseline in the general linear model, $F(1,40) = 13.473$, $p = 0.001$, partial $\eta^2 = 0.252$. In the whole sample (i.e. all tumour type groups combined) valid baseline inspection time scores were positively correlated with the inspection time score at session 2, $r(n = 52) = 0.611$, $p < 0.001$. Similar correlations were obtained for the high-grade glioma group, $r(n = 21) = 0.813$, $p < 0.001$; the metastasis group, $r(n = 8) = 0.886$, $p = 0.003$; and for the 'other' tumour type group, $r(n = 2) = 1.00$, $p < 0.001$. There was no significant correlation between the two scores in the low-grade glioma group, $r(n = 16) = 0.367$, $p = 0.162$, or in the meningioma group, $r(n = 5) = 0.675$, $p = 0.211$. There was no significant main effect of the covariate age, $F(1,40) = 1.077$, $p = 0.306$, partial $\eta^2 = 0.026$. The covariate NART also had no significant main effect in the model, $F(1,40) = 1.051$, $p = 0.312$, partial $\eta^2 = 0.026$. Sex had no significant effect in the model, $F(1,40) = 0.006$, $p = 0.938$, partial $\eta^2 < 0.001$. There was no significant main effect of tumour type in the model that included the effects of the covariates age, NART and baseline valid inspection time score, $F(4,40)$

= 0.506, p = 0.731, partial η^2 = 0.048. The interaction between tumour type and sex was not significant, F(3,40) = 0.629. p = 0.601, partial η^2 = 0.045. That is, in the presence of the effects of age, NART score and baseline test score there was no significant overall difference between the performance of the tumour type groups on inspection time at baseline and session 2 follow-up.

The estimated marginal mean session 2 inspection time score, adjusted for age, NART and baseline inspection time score—derived from the above-described general linear model—was 114.3 (SE 4.3) for the low-grade glioma group; 120.9 (SE 3.9) for the high-grade glioma group; 115.9 (SE 6.1) for the metastasis group; 117.8 (SE 9.7) for the meningioma group and 127.7 (SE 11.6) for the 'other' tumour group. Pairwise comparisons using Least Significant Difference tests revealed no significant differences between the low-grade and high-grade glioma patient groups.

9.2.3 Digit Symbol Coding

Fifteen low-grade glioma patients, 25 high-grade glioma patients, 8 metastasis, 9 meningioma and 3 patients with 'other' tumours all completed the digit symbol coding measure at both baseline and session 2.

The mean score for the low-grade glioma group at baseline was 70.9 (SD 12.2) and at session 2 was 64.1 (SD 14.2). The high-grade glioma group scored a mean of 59.4 (SD 20.8) at baseline and 58.4 (SD 19.6) at session 2. The mean baseline score for the metastasis group was 62.6 (SD 23.7) and at session 2 was 54.5 (SD 26.7) and the meningioma group baseline mean was 48.2 (SD 24.6) and at session 2 was 42.4 (SD 22.8). The other tumour group baseline mean was 81.3 (SD 26.0) and at session 2 was 84.3 (SD 23.0). The mean (standard error) scores on digit symbol coding at baseline and session 2 for each tumour type group are shown in Figure 9.3.

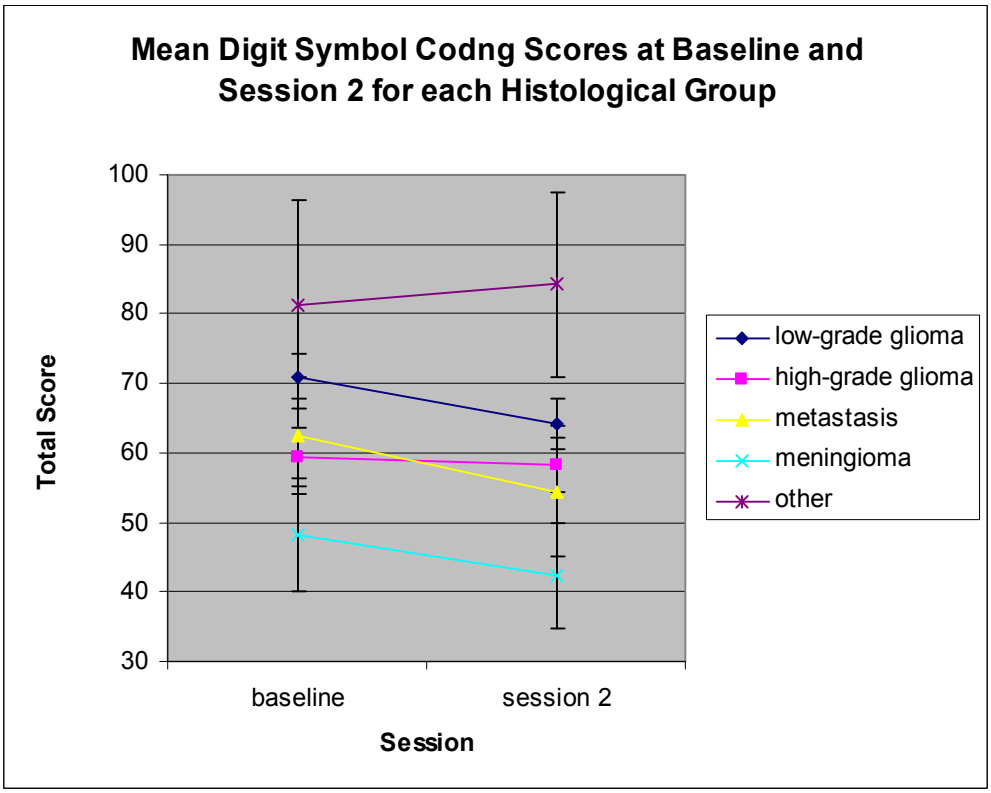


Figure 9.3. Baseline and session 2 digit symbol coding scores for each tumour type group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with digit symbol coding score entered as the dependent variable, revealed a significant main effect of baseline digit symbol coding score, $F(1,48) = 133.269$, $p < 0.001$, partial $\eta^2 = 0.735$. In the whole sample (i.e. the tumour type groups combined) baseline digit symbol coding score was positively correlated with the same test score at session 2, $r(n = 60) = 0.892$, $p < 0.001$. Similar positive correlations were obtained for the high-grade glioma group, $r(n = 25) = 0.965$, $p < 0.001$; the metastasis group, $r(n = 8) = 0.951$, $p < 0.001$; the meningioma group, $r(n = 9) = 0.843$, $p = 0.004$ and for the 'other' tumour group, $r(n = 3) = 0.998$, $p = 0.044$. The low-grade glioma group showed a tendency towards a positive correlation between the two scores but this did not reach a conventional level of statistical significance, $r(n = 15) = 0.497$, $p = 0.060$. There was no significant main effect of the covariate age in the model, $F(1,48) = 0.401$, $p = 0.530$, partial $\eta^2 = 0.008$; or of the covariate NART score, $F(1,48) = 0.666$, $p = 0.419$, partial $\eta^2 = 0.014$. The effect of sex was significant in the

model, $F(1,48) = 4.561$, $p = 0.038$, partial $\eta^2 = 0.087$. Female patients were more likely to have lower scores than male patients. The effect of tumour type was not significant in the model that included the effects of the covariates age, NART score or baseline digit symbol coding score, $F(4,48) = 1.572$, $p = 0.197$, partial $\eta^2 = 0.116$. There was no significant interaction between sex and tumour type, $F(3,48) = 1.377$, $p = 0.261$, partial $\eta^2 = 0.079$. In the presence of the effects of age, NART score and baseline test score there was no significant overall difference between the performance of the tumour type groups on digit symbol coding at baseline and session 2 follow-up.

The estimated marginal mean session 2 digit symbol coding score, adjusted for age, NART and baseline inspection time score—derived from the above-described general linear model—was 57.1 (SE 2.7) for the low-grade glioma group; 60.2 (SE 2.0) for the high-grade glioma group; 53.0 (SE 3.5) for the metastasis group; 64.2 (SE 5.3) for the meningioma group and 68.3 (SE 5.6) for the 'other' tumour group. Pairwise comparisons revealed no significant difference between the low-grade and high-grade glioma tumour groups.

9.2.4 Williams Delayed Recall Test (EFIT)

Fourteen low-grade glioma patients, 26 high-grade glioma patients, 9 with a metastasis, 9 with a meningioma and 3 patients with 'other' tumours who completed the Williams delayed recall test at both baseline and post-operatively during session 2.

The low-grade glioma group scored a mean of 6.4 (SD 3.3) at baseline and 10.9 (SD 7.8) at session 2. The high-grade glioma group had a baseline mean score of 9.2 (SD 4.9) and a session 2 mean score of 11.1 (SD 4.7). The baseline mean score for the metastasis group was 8.9 (SD 6.6) and at session 2 was 12.2 (SD 6.2). The meningioma group scored a baseline mean of 10.1 (SD 7.0) and a mean of 11.4 (SD 7.9) at session 2. The 'other' tumour group baseline mean was 10.3 (SD 4.2) and the session 2 mean score was

14.3 (SD 2.9) for this group. The mean (standard error) scores at baseline and session 2 for each tumour type group are shown in Figure 9.4.

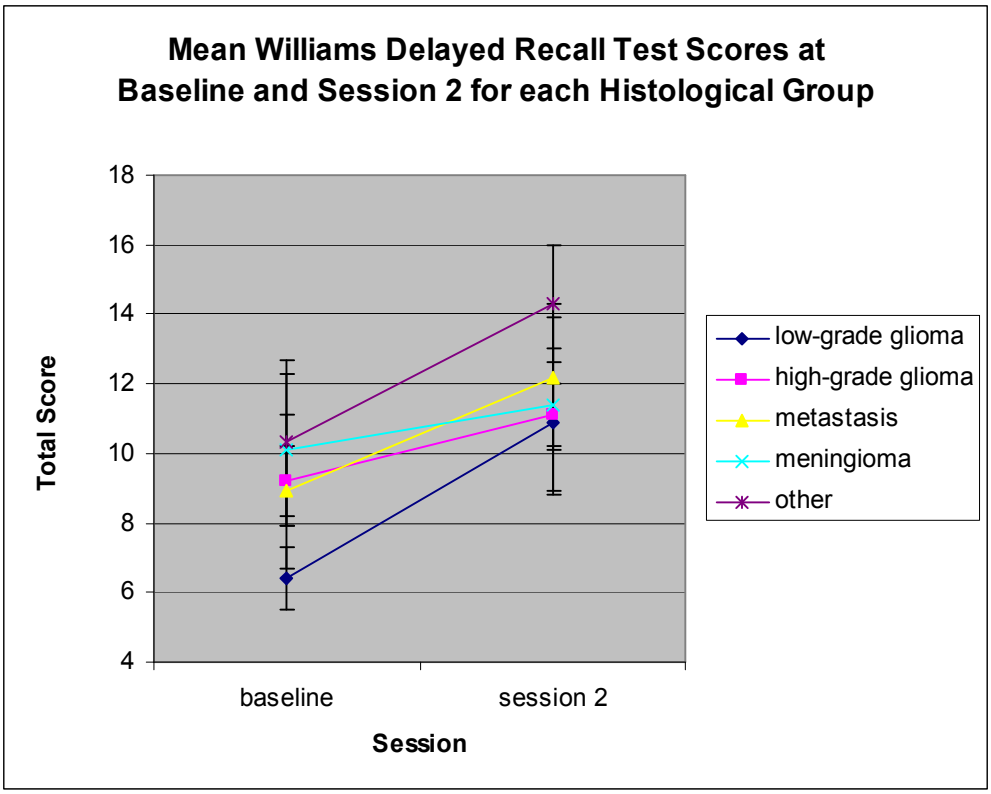


Figure 9.4. Baseline and session 2 Williams Delayed Recall Test scores for each tumour type group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with the session 2 test score entered as the dependent variable revealed a significant main effect of the baseline test score, $F(1,50) = 16.866$, $p < 0.001$, partial $\eta^2 = 0.252$. In the whole sample (i.e. the tumour type groups combined), baseline Williams delayed recall test score was positively correlated with session 2 score, $r(n = 62) = 0.495$, $p < 0.001$. The two scores were also positively correlated in the high-grade glioma group alone, $r(n = 26) = 0.550$, $p = 0.004$, in the metastasis group alone, $r (n = 9) = 0.805$, $p = 0.009$ and in the meningioma group alone, $r(n = 10) = 0.756$, $p = 0.011$. There was no significant correlation between the two test scores in neither the low-grade glioma group, $r(n = 14) = 0.008$, $p = 0.978$; nor in the 'other' tumour group, $r(n = 3) = -0.971$, $p = 0.154$. There was a significant main effect of age in the model, $F(1,50) =$

5.795, p = 0.020, partial $\eta^2 = 0.104$. Older participants tended to have higher (poorer) scores. The effect of the covariate NART score was not significant, F(1,50) = 2.977, p = 0.091, partial $\eta^2 = 0.056$. The effect of sex was not significant in the model, F(1,50) = 0.016, p = 0.901, partial $\eta^2 < 0.001$. The effect of tumour type had no significant main effect in the model that included the effects of the covariates age, NART score and baseline test score, F(4,50) = 1.560, p = 0.199, partial $\eta^2 = 0.111$. The interaction between sex and tumour group was significant in the model, F(3.50) = 4.271, p = 0.009, partial $\eta^2 = 0.204$. Female participants in the low-grade glioma and meningioma group had significantly higher (poorer) scores on this test at session 2 than their male counterparts.

The estimated marginal mean session 2 Williams delayed recall test score, adjusted for age, NART and baseline score—derived from the above-described general linear model—was 14.3 (SE 1.5) for the low-grade glioma group; 10.1 (SE 1.0) for the high-grade glioma group; 11.1 (SE 1.8) for the metastasis group; 10.3 (SE 2.8) for the meningioma group and 14.8 (SE 3.0) for the 'other' tumour group. Pairwise comparisons using Least Significant Difference tests revealed a significant difference between the low-grade glioma and high-grade glioma groups (p = 0.031). This suggests that the low-grade glioma group deteriorated significantly between baseline and session 2 testing by comparison with the high-grade glioma group.

9.2.5 Nine Hole Peg Test (Right Hand, EFIT)

Fifteen patients in the low-grade glioma group, 26 in the high-grade glioma group, 8 with a metastasis, 10 with a meningioma and 3 'other' tumour patients completed the right hand nine hole peg test at both baseline and session 2.

The low-grade glioma group had a mean baseline score of 12.5 (SD 2.0) and a mean of 14.2 (SD 2.0) at session 2. The baseline mean score for the high-grade glioma group was 14.0 (SD 3.0) and at session 2 the group mean was 14.5 (SD 3.1). The metastasis group

had a mean score of 13.3 (SD 2.3) at baseline and 13.7 (SD 3.0) at session 2. The meningioma group had mean scores of 16.0 (SD 3.6) and 16.2 (SD 2.8) at baseline and session 2, respectively. The mean baseline score for the 'other' tumour group was 12.4 (SD 0.2) and at session 2 was 12.3 (SD 0.4). The mean (standard error) scores for each tumour type group at baseline and session 2 are shown in Figure 9.5.

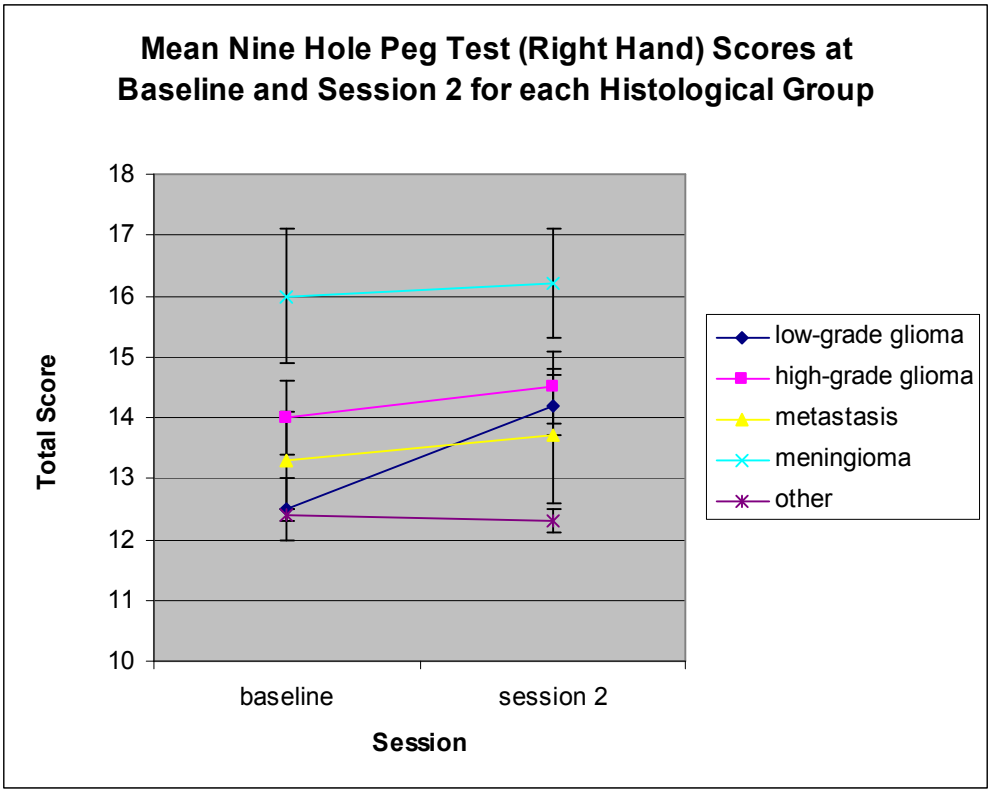


Figure 9.5. Baseline and session 2 right hand nine hole peg test scores for each tumour type group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with right hand nine hole peg test score at session 2 entered as the dependent variable revealed a significant main effect of the test score at baseline, $F(1,50) = 35.156$, $p < 0.001$, partial $\eta^2 = 0.413$. In the whole sample (i.e. all tumour type groups combined) baseline right hand nine hole peg test score was correlated with the corresponding score at session 2, $r(n = 62) = 0.741$, $p < 0.001$. Similar correlations were obtained for the two scores in the low-grade glioma group alone, $r(n = 15) = 0.673$, $p = 0.006$; the high-grade glioma group, $r(n = 26) = 0.697$, $p < 0.001$; the metastasis group,

r(n = 8) = 0.718, p = 0.045; the meningioma group, r(n = 10) = 0.872, p = 0.001. The two scores were, however, not significantly correlated in the 'other' tumour group alone, r(n = 3) = -0.823, p = 0.384. There was no significant effect of either of the covariates age, F(1,50) = 1.271, p = 0.265, partial $\eta^2 = 0.025$; or NART score, F(1,50) = 0.024, p = 0.877, partial $\eta^2 < 0.001$. The effect of sex was not significant, F(1,50) = 0.025, p = 0.876, partial $\eta^2 < 0.001$. The effect of tumour type was not significant in the model that included the effects of the covariates age, NART score and baseline test score, F(4,50) = 0.622, p = 0.649, partial $\eta^2 = 0.047$. There was no significant interaction between sex and tumour type, F(3,50) = 1.843, p = 0.151, partial $\eta^2 = 0.100$.

The estimated marginal mean session 2 right hand nine hole peg test score, adjusted for age, NART and corresponding baseline test score—derived from the above-described general linear model—was 14.8 (SE 0.6) for the low-grade glioma group; 14.5 (SE 0.4) for the high-grade glioma group; 14.3 (SE 0.7) for the metastasis group; 13.8 (SE 1.1) for the meningioma group and 13.1 (SE 1.1) for the 'other' tumour group. Pairwise comparisons using Least Significant Difference tests revealed no significant differences between the low-grade and high-grade glioma tumour groups.

9.2.6 Nine Hole Peg Test (Left Hand, EFIT)

Fourteen patients in the low-grade glioma group, 26 in the high-grade glioma group, 8 patients with a metastasis, 10 with a meningioma and 3 with 'other' tumour types completed the left hand nine hole peg test at both baseline and session 2.

The mean baseline score for the low-grade glioma group on the left hand nine hole peg test was 13.3 (SD 1.7) and at session 2 was 14.5 (SD 2.4). The high-grade glioma group had a mean score of 16.0 (SD 3.5) at baseline and 15.9 (SD 3.5) at session 2. The metastasis group had mean scores of 13.8 (SD 1.7) and 15.4 (SD 1.5) at baseline and session 2, respectively. The mean baseline score for the meningioma group was 18.5 (SD 4.6) and at session 2 was 18.7 (SD 4.4). The 'other' tumour type group had a mean

of 13.1 (SD 2.2) at baseline and 13.3 (SD 2.0) at session 2. The mean (standard error) scores for each tumour type group at baseline and session 2 on the left hand nine hole peg test are shown in Figure 9.6.

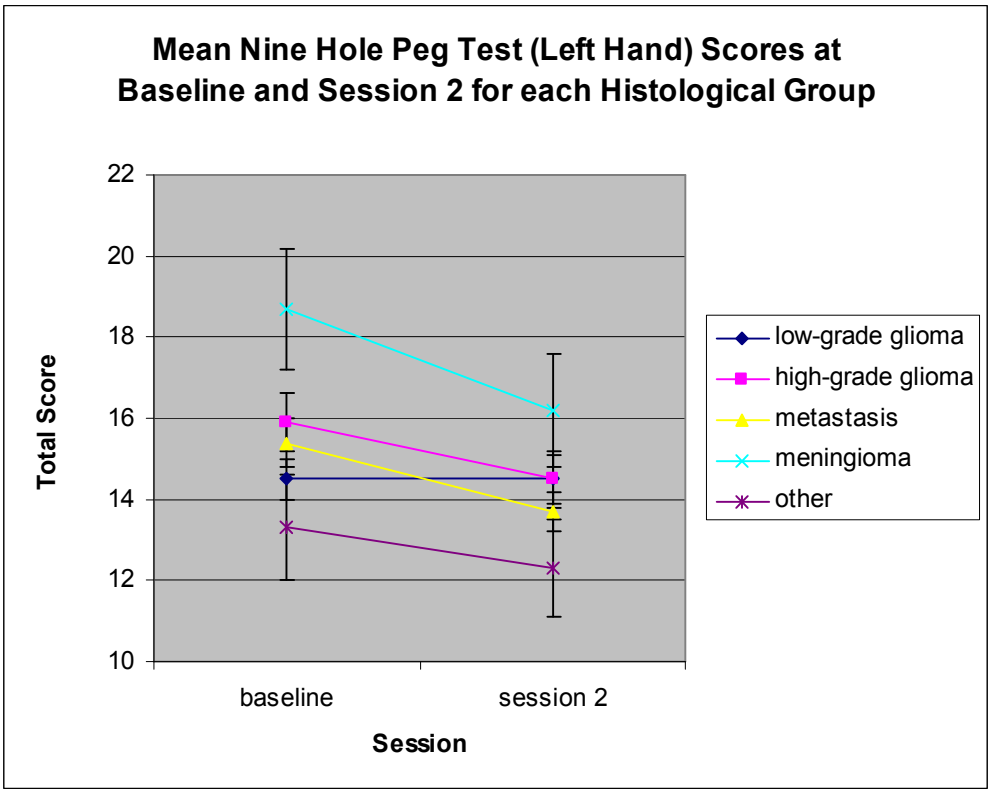


Figure 9.6. Baseline and session 2 left hand nine hole peg test scores for each tumour type group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with left hand nine hole peg test score entered as the dependent variable revealed a significant main effect of the corresponding baseline score, $F(1,49) = 99.667$, $p < 0.001$, partial $\eta^2 = 0.670$. In the whole sample of all tumour type groups combined, baseline left hand nine hole peg test score was positively correlated with session 2 left hand nine hole peg test score, $r(n = 61) = 0.859$, $p < 0.001$. Similar correlations were obtained for the low-grade glioma group, $r(n = 14) = 0.696$, $p = 0.006$; high-grade glioma group, $r(n = 26) = 0.814$, $p < 0.001$; metastasis group, $r(n = 8) = 0.886$, $p = 0.003$ and meningioma group, $r(n = 10) = 0.954$, $p < 0.001$. The two scores were not significantly correlated in the 'other' tumour type group, $r(n = 3) =$

0.658, p = 0.543. There was a significant main effect of the covariate age n the model, F(1,49) = 6.642, p = 0.013, partial η^2 = 0.119. Older patients took significantly longer to complete the task. The covariate NART score had no significant effect in the model, F(1,49) = 0.812, p = 0.372, partial η^2 = 0.016. There was also no significant main effect of sex, F(1,49) = 1.344, p = 0.252, partial η^2 = 0.027. The effect of tumour type was not significant in the model that included the effects of the covariates age, NART score and baseline left hand nine hole peg test score, F(4,49) = 1.098, p = 0.368, partial η^2 = 0.082. The interaction between sex and tumour type was non-significant, F(3,49) = 0.326, p = 0.807, partial η^2 = 0.020. That is, in the presence of the effects of age, NART score and corresponding baseline test score, there was no significant overall difference between the performance of the tumour type subgroups on the left hand nine hole peg test at session 2 post-operative follow-up.

The estimated marginal mean session 2 scores on the left hand nine hole peg test, adjusted for age, NART score and baseline left hand nine hole peg test score – derived from the above-described general linear model – were 16.4 (SE 0.5) for the low-grade glioma group; 15.3 (SE 0.3) for the high-grade glioma group; 16.3 (SE 0.7) for the metastasis group; 16.3 (SE 1.0) for the meningioma group and 15.6 (SE 1.0) for the 'other' tumour type group. Pairwise comparisons, using LSD tests, revealed no significant differences between the two glioma patient groups (low-grade and high-grade).

9.2.7 Timed Ten Metre Walk (EFIT)

Fourteen patients in the low-grade glioma group, 25 in the high-grade glioma group, 9 patients with a metastasis, 5 with a meningioma and 3 with 'other' tumour types completed the timed ten metre walk test at both baseline and session 2.

The mean baseline score for the low-grade glioma group was 6.0 (SD 1.0) and at session 2 was 6.4 (SD 1.0). The high-grade glioma group had a mean score of 6.8 (SD 1.1) at

baseline and 7.1 (SD 1.3) at session 2. The metastasis group scored a mean of 7.5 (SD 1.8) at baseline and 8.5 (SD 2.7) at session 2. The meningioma group scored 7.9 (SD 1.8) and 7.2 (SD 1.6) at baseline and session 2, respectively. The 'other' tumour type group scored a mean of 5.2 (SD 1.0) at baseline and 6.3 (SD 0.4) at session 2. The mean (standard error) scores at baseline and session 2 for each tumour type group are shown in Figure 9.7.

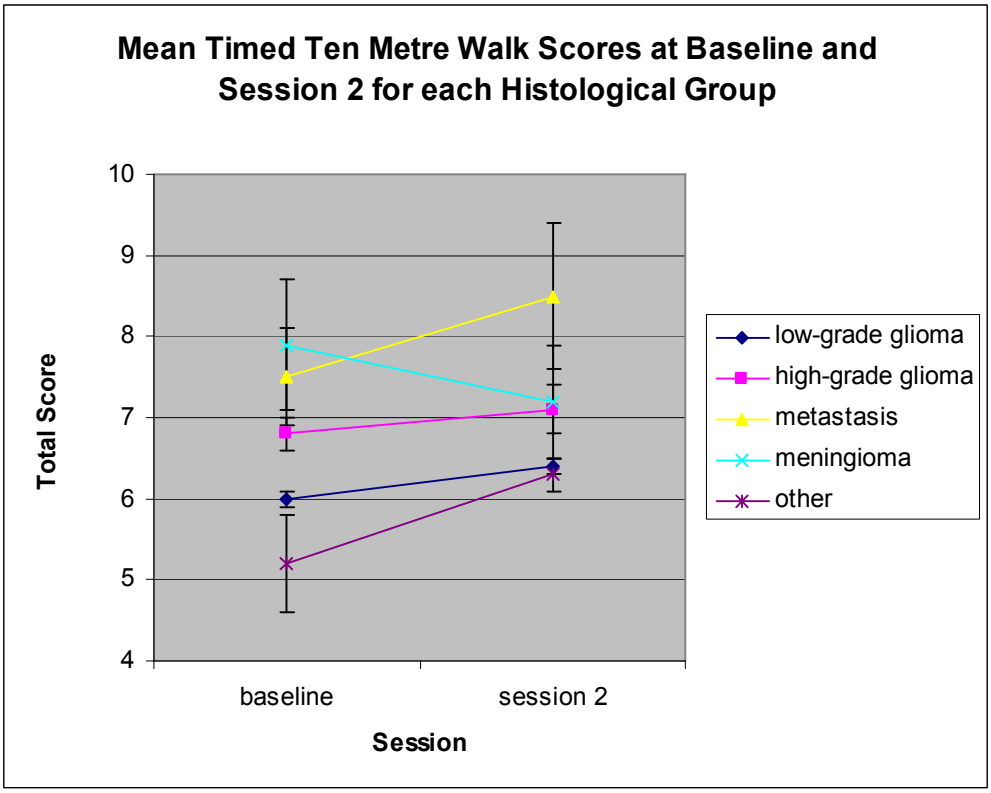


Figure 9.7. Baseline and session 2 timed ten metre walk test scores for each tumour type group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with session 2 timed ten metre walk test score entered as the dependent variable revealed a significant main effect of the baseline timed ten metre walk score, $F(1,44) = 34.545$, $p < 0.001$, partial $\eta^2 = 0.440$. In the whole sample of all tumour type groups combined, baseline timed ten metre walk score was positively correlated with the same score at session 2, $r(n = 56) = 0.731$, $p < 0.001$. The two scores were similarly correlated in the low-grade glioma group, $r(n = 14) = 0.664$, $p = 0.010$;

the high-grade glioma group, r(n = 25) = 0.549, p = 0.004 and the metastasis group, r(n = 9) = 0.924, p < 0.001. The two scores were not significantly correlated in the meningioma group, r(n = 5) = 0.760, p = 0.136 or in the 'other' tumour type group, r(n = 3) = 0.702, p = 0.505. There was no significant main effect of the covariate age, F(1,44) = 1.043, p = 0.313, partial η^2 = 0.023, or the covariate NART score, F(1,44) = 0.032, p = 0.860, partial η^2 = 0.001. Sex had no significant effect in the model, F(1,44) = 0.416, p = 0.522, partial η^2 = 0.009. There was no significant main effect of tumour type in the model that included the effects of the covariates age, NART score and baseline timed ten metre walk score, F(4,44) = 0.975, p = 0.431, partial η^2 = 0.081. The interaction between tumour type and sex was not significant, F(3,44) = 1.163, p = 0.334, partial η^2 = 0.073. That is, in the presence of the effects of age, NART score and baseline test score, there was no significant overall difference between the performance of the tumour type subgroups on the timed ten metre walk at session 2, post-operatively.

The estimated marginal mean session 2 timed ten metre walk scores, adjusted for age, NART score and baseline timed ten metre walk score – derived from the above-described general linear model – were 7.2 (SE 0.3) for the low-grade glioma group; 7.0 (SE 0.2) for the high-grade glioma group; 7.8 (SE 0.4) for the metastasis group; 6.5 (SE 0.7) for the meningioma group and 7.6 (SE 0.7) for the 'other' tumour type group. Pairwise comparisons using LSD tests revealed no significant difference between the low-grade and high-grade glioma groups.

Table 9.1. Overview of comparisons of post-operative test performance in each histological group

Test	Effect of tumour type		
	F	Sig. (p)	n^{2*}
Inspection Time (All Data)	1.100	0.368	0.084
Inspection Time (Valid Data)	0.506	0.731	0.048
Digit Symbol Coding (Total)	1.572	0.197	0.116
EFIT Williams Delayed Recall Test (total)	1.560	0.199	0.111
EFIT Nine Hole Peg Test (Right Hand, secs)	0.622	0.649	0.047
EFIT Nine Hole Peg Test (Left Hand, secs)	1.098	0.360	0.082
EFIT Timed Ten Metre Walk (secs)	0.975	0.431	0.081

*n^2 = the proportion of variance accounted for by the covariate (tumour lobe).

9.3 Tumour Lobe

Of the total cohort of brain tumour patients who were tested at baseline and also post-operatively at session 2, 21 patients had a tumour located in the frontal lobes, 10 had a temporal lobe tumour, 2 had a limbic tumour, 9 had a parietal lobe tumour, 4 had an occipital lobe tumour, 12 patients had tumours that infiltrated more than one lobe of the brain ('multiple lobes') and 6 patients had tumours located elsewhere (e.g. pituitary gland).

9.3.1 Inspection Time Scores: All Inspection Time Data

Of the cohort of brain tumour patients who completed both baseline and session 2 testing, 3 patients with tumours located in 'multiple lobes' and 1 patient with a tumour located elsewhere ('other') did not complete inspection time testing. Therefore, data from 21 patients with frontal tumours, 10 with temporal lobe tumours, 2 with limbic tumours, 9 with parietal lobe tumours, 4 with occipital lobe tumours, 9 with tumours in multiple lobes and 5 with 'other' tumours were included in the analysis.

The mean baseline inspection time score for the frontal lobe group was 119.0 (SD 21.0) and at session 2 was 117.5 (SD 20.2). The temporal lobe group scored a baseline mean of 124.2 (SD 11.4) and a session 2 mean score of 119.6 (SD 14.8). The limbic group scored 119.5 (SD 20.5) and 110.5 (SD 27.6) at baseline and session2, respectively. The baseline mean for the parietal lobe group was 115.8 (SD 22.3) and the session 2 mean was 107.9 (SD 21.0). The group of patients with occipital lobe tumours had a baseline mean score of 78.0 (SD 25.3) and a session 2 mean score of 89.3 (SD 34.7). The multiple lobe tumour group had a baseline mean of 120.9 (SD 12.0) and a session 2 mean of 110.9 (SD 29.1) and the group of patients with tumours located elsewhere ('other') had mean scores of 127.4 (SD 9.1) and 125.0 (SD15.9) at baseline and session

2, respectively. The mean (standard error) scores at baseline and session 2 for each tumour group are shown in Figure 9.8.

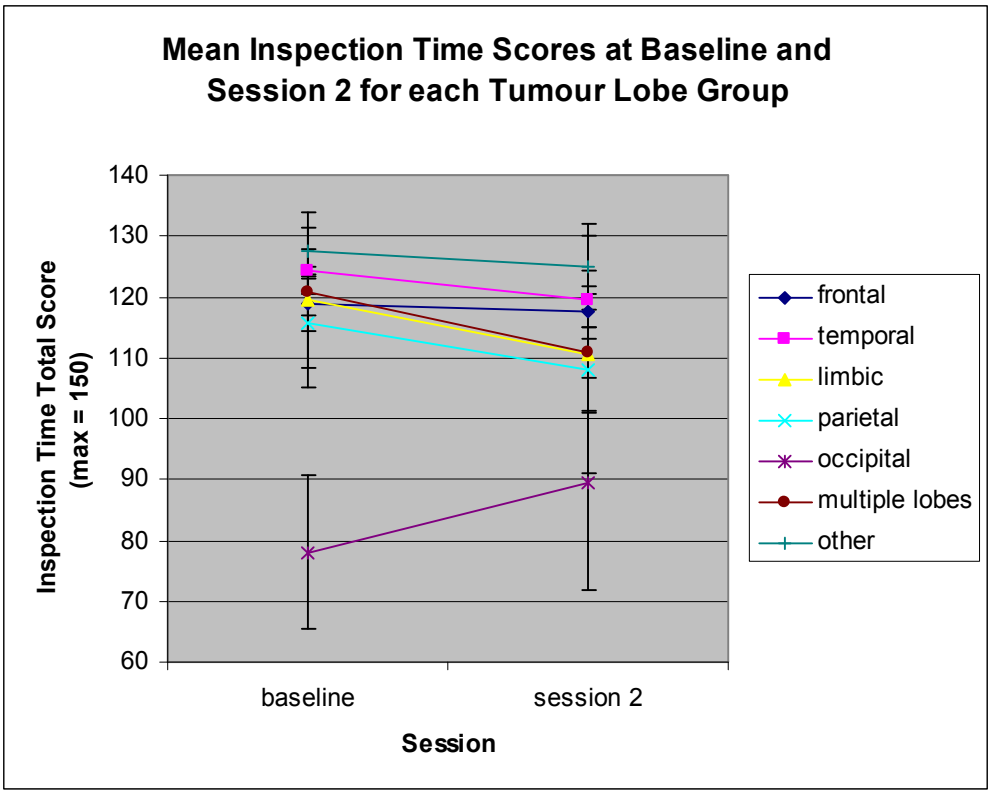


Figure 9.8. Baseline and session 2 inspection time scores (all data) for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

Post-operative (session 2) inspection time score was entered as the dependent variable in the general linear model. There was a significant main effect of baseline inspection time score in the model, $F(1,44) = 15.783$, $p < 0.001$, partial $\eta^2 = 0.264$. Participants with higher baseline inspection time scores were significantly more likely to have higher scores at session 2. The effect of the covariate age did not reach significance in the model, $F(1,44) = 3.404$, $p = 0.072$, partial $\eta^2 = 0.072$. NART score also had no significant main effect, $F(1,44) = 0.953$, $p = 0.334$, partial $\eta^2 = 0.021$. The effect of sex was not significant in the model, $F(1,44) = 0.081$, $p = 0.778$, partial $\eta^2 = 0.002$. There was no significant main effect of tumour lobe in the model that included the effects of the covariates age, NART score and baseline inspection time score, $F(6,44) = 0.495$, $p =$

0.808, partial $\eta^2 = 0.063$. The interaction between tumour lobe and sex was not significant, $F(5,44) = 0.471$, $p = 0.796$, partial $\eta^2 = 0.051$.

The estimated marginal mean session 2 inspection time score, adjusted for age, NART score and baseline inspection time score – derived from the above-described general linear model – were 117.0 (SE 3.7) for the frontal lobe group; 114.3 (SE 5.8) for the temporal lobe group; 103.6 (SE 12.2) for the limbic group; 110.1 (SE 6.0) for the parietal lobe group; 118.1 (SE 10.0) for the occipital lobe group; 108.5 (SE 5.6) for the multiple lobe group and 116.5 (SE 7.8) for the 'other' location group.

9.3.2 Inspection Time Scores: Valid Inspection Time Data

When data from only patients with 'valid' baseline inspection time scores were included in the analysis, there were 18 frontal lobe tumour patients, 10 temporal, 2 limbic, 8 parietal, 9 'multiple lobe' and 5 'other' location patients who also had inspection time scores at session 2.

The mean score at baseline for the frontal lobe group was 125.5 (SD 13.5) and at session 2 was 121.3 (SD 17.2). The temporal lobe group had a baseline mean of 124.2 (SD 11.4) and a mean at session 2 of 119.6 (SD 14.8). The mean scores for the limbic group were 119.5 (SD 20.5) and 110.5 (SD 27.6) at baseline and session 2, respectively. The parietal lobe group had a baseline mean of 121.4 (SD 15.7) and a session 2 mean score of 110.8 (SD 20.5). The mean scores for the multiple lobe group were 120.9 (SD 12.0) and 110.9 (SD 29.1) for baseline and session 2, respectively. The baseline mean for the 'other' location group was 127.4 (SD 9.1) and was125.0 (SD 15.9) at session 2. The occipital group who returned for session2 follow-up all had invalid baseline inspection time scores. The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.9.

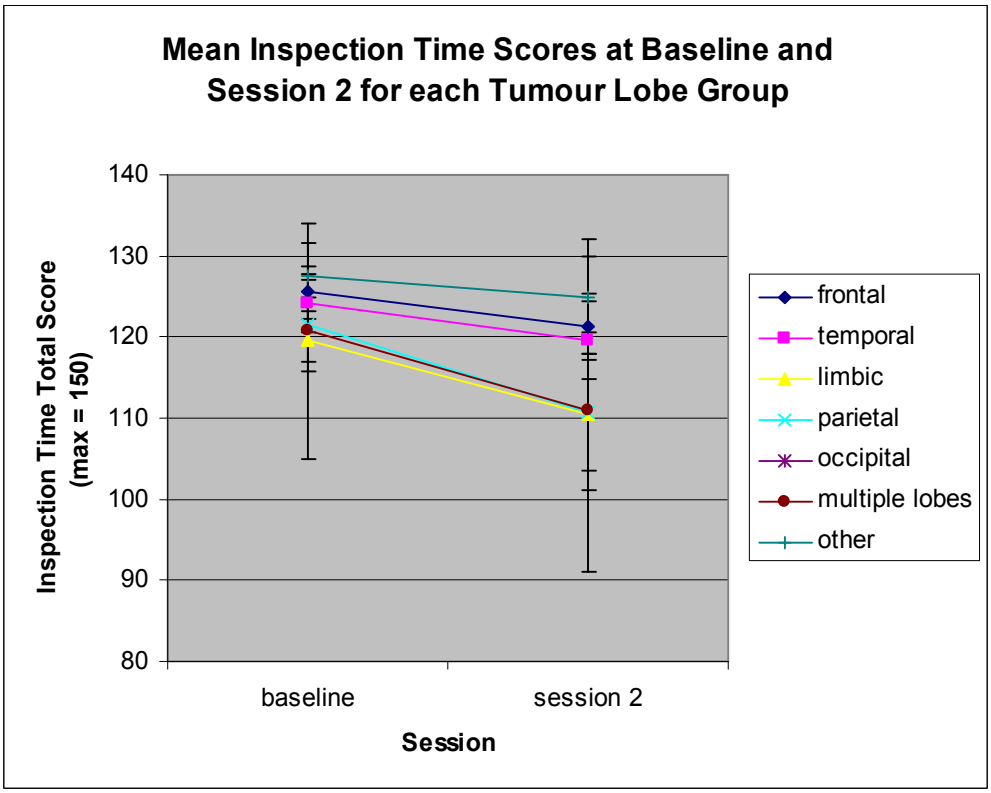


Figure 9.9. Baseline and session 2 inspection time scores (valid baseline data) for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

Session 2 inspection time score was entered as the dependent variable in the general linear model. There was a significant main effect of the baseline inspection time score, $F(1,38) = 5.105$, $p = 0.030$, partial $\eta^2 = 0.118$. Patients with higher scores at baseline were significantly more likely to have higher scores at session 2. The effects of the covariate age did not reach statistical significance, $F(1,38) = 3.563$, $p = 0.067$, partial $\eta^2 = 0.086$. The covariate NART score also had no significant main effect, $F(1,38) = 3.200$, $p = 0.082$, partial $\eta^2 = 0.078$. Sex also had no significant effect in the model, $F(1,38) < 0.001$, $p = 0.993$, partial $\eta^2 < 0.001$. The effect of tumour lobe was not significant in the model that included the effects of the covariates age, NART score and baseline inspection time score, $F(5,38) = 0.759$, $p = 0.585$, partial $\eta^2 = 0.091$. There was no significant interaction between sex and tumour lobe, $F(4,38) = 0.514$, $p = 0.726$, partial $\eta^2 = 0.051$.

The estimated marginal mean scores, adjusted for age, NART score and baseline inspection time score – derived from the above-described general linear model – were 121.8 (SE 4.0) for the frontal lobe group; 118.9 (SE 5.6) for the temporal lobe group; 106.0 (SE 12.2) for the limbic group; 111.2 (SE 6.2) for the parietal group; 113.4 (SE 5.5) for the multiple lobe group and 122.5 (SE 7.6) for the 'other' location group.

9.3.3 Digit Symbol Coding

There were 21 frontal lobe patients, 10 temporal lobe patients, 2 patients with limbic tumours, 7 parietal lobe patients, 3 occipital, 11 multiple lobe and 6 patients with tumours located elsewhere ('other') who completed digit symbol coding at both baseline and session 2, post-operatively.

The mean baseline digit symbol coding score for the frontal lobe group was 59.8 (SD 16.7) and at session 2 was 58.8 (SD 18.2). The temporal lobe group had a baseline mean score of 70.7 (SD 21.5) and a session2 mean score of 64.3 (SD 24.6). The limbic group scored a mean of 72.0 (SD 7.1) at baseline and 64.0 (SD 12.7) at session 2. The mean scores for the parietal group were 54.0 (SD 21.4) and 48.1 (SD 17.6) at baseline and session 2, respectively. The occipital lobe tumour group had a baseline mean digit symbol coding score of 38.0 (SD 30.3) and a mean of 36.3 (SD 30.7) at session 2. The mean baseline score for the multiple lobe group was 65.4 (SD 22.5) and at session 2 was 59.2 (SD 19.9). The 'other' location group score means of 68.7 (SD 27.2) and 65.3 (SD 29.4) at baseline and session 2, respectively. The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.10.

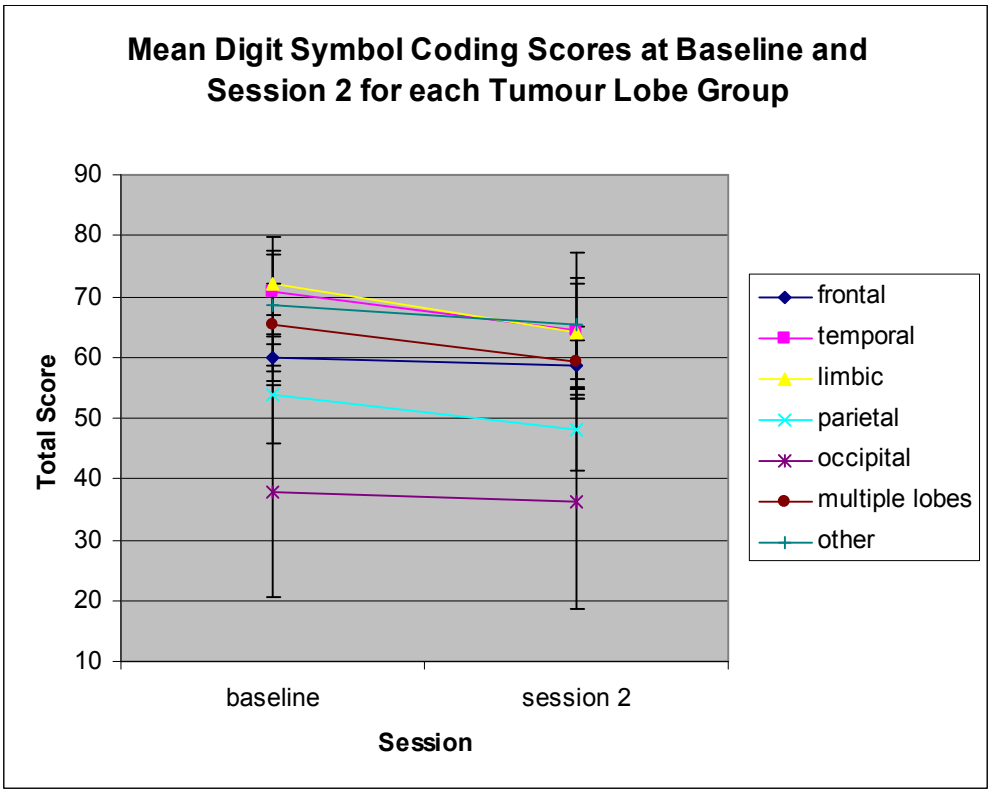


Figure 9.10. Baseline and session 2 digit symbol coding scores for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with session 2 digit symbol coding score as the dependent variable revealed a significant main effect of the corresponding baseline test score, $F(1,44) = 110.451$, $p < 0.001$, partial $\eta^2 = 0.715$. There was no significant effect of either covariate age, $F(1,44) = 0.018$, $p = 0.893$, partial $\eta^2 < 0.001$; or NART score, $F(1,44) = 0.036$, $p = 0.851$, partial $\eta^2 = 0.001$. Sex had a significant main effect in the model, $F(1,44) = 7.067$, $p = 0.011$, partial $\eta^2 = 0.138$. Female participants had significantly lower session 2 scores than their male counterparts. There was no significant main effect of tumour lobe in the model that included the effects of the covariates age, NART score and baseline digit symbol coding score, $F(6,44) = 0.792$, $p = 0.581$, partial $\eta^2 = 0.097$. The interaction between sex and tumour lobe was not significant, $F(5,44) = 1.753$, $p = 0.143$, partial $\eta^2 = 0.166$.

The estimated marginal mean scores, adjusted for age, NART score and baseline test score – derived from the above-described general linear model – were 60.7 (SE 2.1) for the frontal lobe group; 54.7 (SE 3.3) for the temporal lobe group; 55.5 (SE 6.9) for the limbic group; 55.9 (SE 3.7) for the parietal group; 55.3 (SE 6.2) for the occipital group; 56.7 (SE 2.9) for the multiple lobe group and 59.9 (SE 3.9) for the 'other' location group.

9.3.4 Williams Delayed Recall Test (EFIT)

Twenty-one patients in the frontal lobe group, 10 in the temporal lobe group. 2 in the limbic group, 8 with parietal tumours, 4 with occipital lobe tumours, 11 with tumours in multiple lobes and 6 patients with tumours elsewhere ('other') completed the Williams delayed recall test at both baseline and session 2.

The mean baseline Williams delayed recall test score for the frontal lobe tumour group was 8.5 (SD 6.2) and at session 2, the mean score was 10.1 (SD 5.4). The temporal lobe group scored a mean of 7.9 (SD 2.6) at baseline and 13.2 (SD 7.5) at session 2. The limbic group baseline mean score was 7.0 (SD 4.2) and at session 2 was 5.5 (SD 2.1). The occipital lobe tumour group mean scores were 17.0 (SD 3.9) and 18.0 (SD 8.1) at baseline and session 2, respectively. The multiple lobe tumour group had a baseline mean score of 7.8 (SD 5.3) and a session 2 mean score of 13.9 (SD 5.3). The 'other' location group mean score was 7.0 (SD 3.2) at baseline and 9.2 (SD 5.4) at session 2. The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.11.

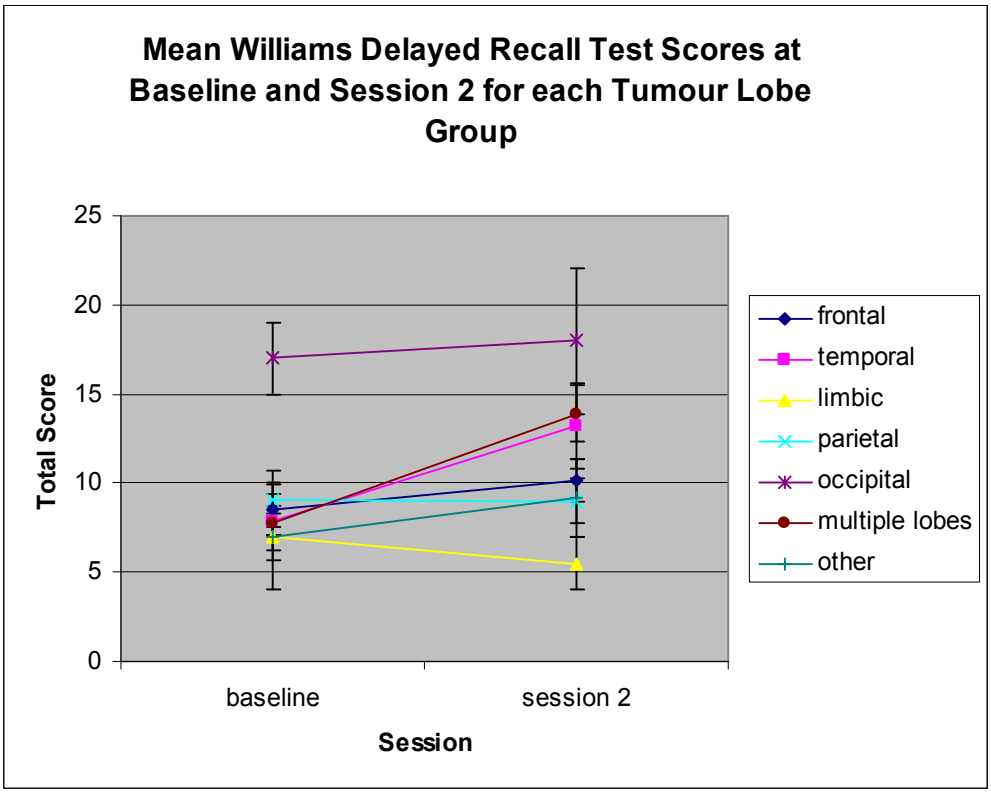


Figure 9.11. Baseline and session 2 Williams Delayed Recall Test scores for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

General linear modelling, with session 2 test score specified as the dependent variable revealed a significant main effect of the Williams delayed recall test score at baseline, $F(1,46) = 8.455$, $p = 0.006$, partial $\eta^2 = 0.155$. Patients with lower (better) scores at baseline were significantly more likely to have lower scores at session 2 testing. There was no significant effect of either covariate age, $F(1,46) = 1.261$, $p = 0.267$, partial $\eta^2 = 0.027$; or NART score, $F(1,46) = 0.060$, $p = 0.808$, partial $\eta^2 = 0.001$. Sex had no significant main effect in the model, $F(1,46) = 0.129$, $p = 0.721$, partial $\eta^2 = 0.003$. The effect of tumour lobe was not significant in the model that included the effects of age, NART score and baseline test score, $F(6,46) = 2.091$, $p = 0.073$, partial $\eta^2 = 0.214$. There was no significant interaction between sex and tumour lobe location, $F(5,46) = 1.942$, $p = 0.106$, partial $\eta^2 = 0.174$.

The estimated marginal mean scores, adjusted for age, NART score and baseline Williams delayed recall test score – derived from the above-described general linear model – were 10.2 (SE 1.1) for the frontal lobe tumour group; 15.0 (SE 1.7) for the temporal lobe tumour group; 7.1 (SE 3.6) for the limbic group; 9.0 (SE 1.9) for the parietal lobe tumour group; 13.8 (SE 2.8) for the occipital lobe tumour group; 13.8 (SE 1.6) for the multiple lobes group and 10.1 (SE 2.1) for the 'other' location group.

9.3.5 Nine Hole Peg Test (Right Hand, EFIT)

Twenty-one patients in the frontal lobe tumour group, 10 in the temporal lobe tumour group, 2 in the limbic group, 7 in the parietal lobe group, 4 in the occipital lobe group, 12 in the multiple lobe group and 6 with tumours elsewhere in the brain ('other') completed this test at both baseline and session 2.

The mean score for the frontal lobe tumour group at baseline was 13.8 (SD 2.2) and at session 2 was 14.5 (SD 2.5). The temporal lobe tumour group had a mean score of 13.2 (SD 2.8) at baseline and 13.7 (SD 2.5) at session 2. The limbic tumour group had mean scores of 13.0 (SD 0.2) and 15.7 (SD 0.5) at baseline and session 2, respectively. The parietal lobe tumour group had a baseline mean score of 15.2 (SD 3.9) and a session 2 mean of 14.9 (SD 2.5). The mean scores for the occipital lobe tumour group were 15.5 (SD 4.7) and 15.1 (SD 3.8) at baseline and session 2, respectively. The multiple lobe tumour group scored a mean of 13.5 (SD 3.7) at baseline and 14.9 (SD 4.2) at session 2 and the 'other' location group had a mean of 13.3 (SD 1.7) at baseline and 13.7 (SD 1.5) at session 2. The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.12.

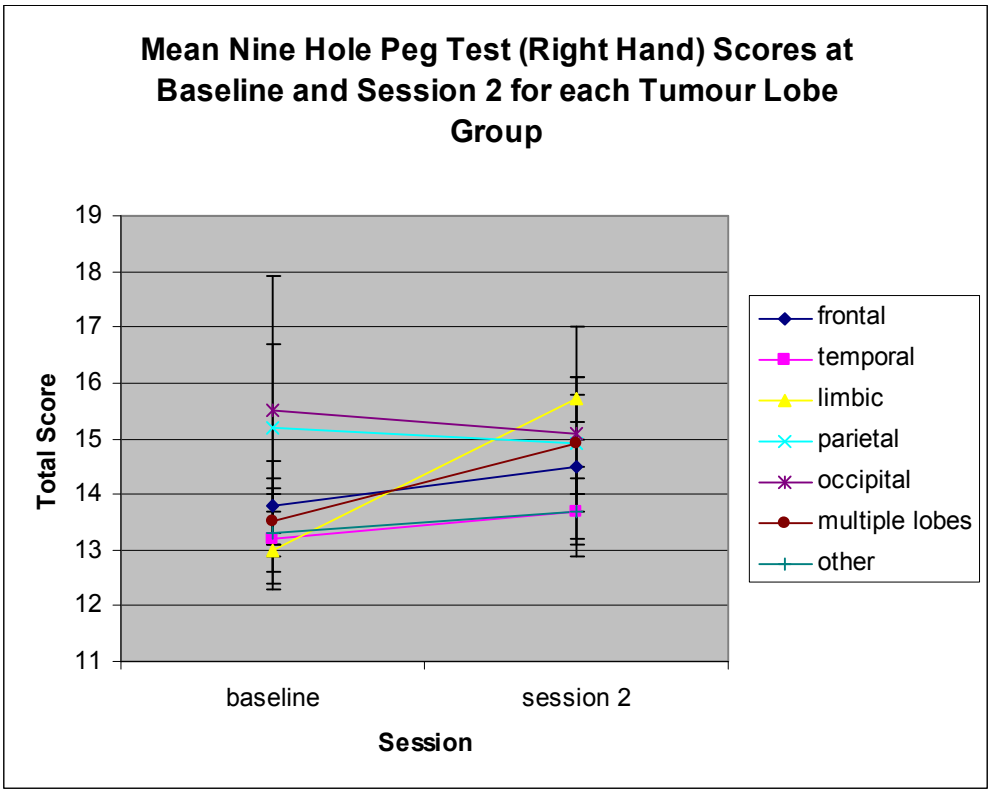


Figure 9.12. Baseline and session 2 nine hole peg test (right hand) scores for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of baseline score on session 2 nine hole peg test (right hand) scores, $F(1,46) = 43.448$, $p < 0.001$, partial $\eta^2 = 0.486$. Those patients who performed more quickly at baseline were significantly more likely to perform more quickly at session 2 on this measure. There was no significant main effect of the covariate age, $F(1,46) = 1.252$, $p = 0.269$, partial $\eta^2 = 0.027$. The covariate NART also had no significant main effect in the model, $F(1,46) = 0.024$, $p = 0.878$, partial $\eta^2 = 0.001$. Sex had no significant main effect in the model, $F(1,46) = 0.011$, $p = 0.918$, partial $\eta^2 < 0.001$. Tumour lobe had no significant main effect in the model that included the effects of the covariates age, NART score and baseline right hand nine hole peg test score, $F(6,46) = 0.641$, $p = 0.697$, partial $\eta^2 = 0.077$. There was no significant interaction between sex and tumour lobe, $F(5,46) = 0.815$, $p = 0.545$, partial $\eta^2 = 0.081$.

The estimated marginal mean scores, adjusted for age, NART score and baseline test score – derived from the above described general linear model – were 14.6 (SE 0.4) for the frontal lobe tumour group; 14.5 (SE 0.7) for the temporal lobe tumour group; 16.0 (SE 1.5) for the limbic group; 13.9 (SE 0.8) for the parietal lobe tumour group; 13.9 (SE 1.0) for the occipital lobe tumour group; 15.2 (SE 0.6) for the multiple lobe tumour group and 14.1 (SE 0.8) for the 'other' location tumour group

9.3.6 Nine Hole Peg Test (Left Hand, EFIT)

Twenty patients in the frontal lobe group, 10 in the temporal lobe group, 2 in the limbic group, 4 in the occipital lobe group, 12 in the multiple lobe group and 6 in the 'other' group all completed this test at both baseline and session 2 follow-up.

The mean baseline score for the frontal lobe tumour group was 15.5 (SD 3.7) and at session 2 was 15.7 (SD 3.5). The temporal lobe tumour group had a baseline mean score of 13.9 (SD 2.7) and a mean score of 14.6 (SD 2.0) at session 2. The limbic tumour group had mean scores of 12.8 (SD 1.3) and 14.7 (SD 0.3) at baseline and session 2, respectively. The parietal lobe tumour group had a mean baseline score of 16.0 (SD 3.0) and a session 2 mean score of 16.2 (SD 3.0). The mean scores for the occipital lobe tumour group were 18.3 (SD 1.9) and 20.1 (SD 3.6) at baseline and session 2, respectively. The group of patients with multiple lobe tumours scored a mean of 15.8 (SD 4.8) at baseline and 16.3 (SD 4.4) at session 2. The 'other' tumour location group scored a mean of 14.4 (SD 3.1) at baseline and at session 2 had a mean score of 14.8 (SD 2.7). The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.13.

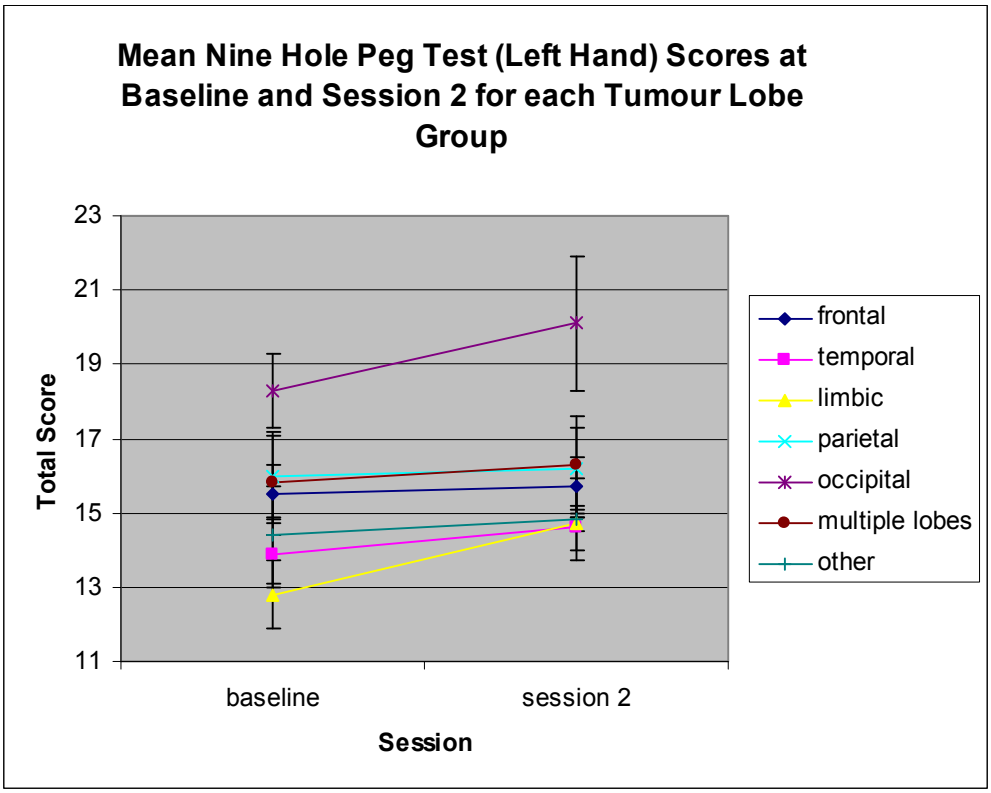


Figure 9.13. Baseline and session 2 nine hole peg test (left hand) scores for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of baseline test score on left hand nine hole peg test scores, $F(1,45) = 92.662$, $p < 0.001$, partial $\eta^2 = 0.673$. Those participants who had faster scores on the test at baseline were significantly more likely to have faster scores at session 2. The covariate age also had a significant main effect in the model, $F(1,45) = 6.296$, $p = 0.016$, partial $\eta^2 = 0.123$. Older participants were significantly more likely to take longer to complete the task. NART score had no significant main effect in the model, $F(1,45) = 0.697$, $p = 0.408$, partial $\eta^2 = 0.015$. Sex also had no significant main effect in the model, $F(1,45) = 0.168$, $p = 0.684$, partial $\eta^2 = 0.004$. Tumour lobe had no significant main effect in the model that included the effects of the covariates age, NART score and baseline test score, $F(6,45) = 0.919$, $p = 0.491$, partial $\eta^2 = 0.109$. The interaction between sex and tumour lobe was not significant, $F(5,45) = 0.105$, $p = 0.991$, partial $\eta^2 = 0.012$.

The estimated marginal mean scores, adjusted for age, NART score and baseline test score – derived from the above described general linear model – were 15.6 (SE 0.4) for the frontal lobe tumour group; 15.8 (SE 0.6) for the temporal lobe tumour group; 17.3 (SE 1.3) for the limbic tumour group; 15.6 (SE 0.7) for the parietal lobe tumour group; 17.6 (SE 0.9) for the occipital lobe tumour group; 15.8 (SE 0.5) for the multiple lobes tumour group and 15.6 (SE 0.7) for the 'other' location group.

9.3.7 Timed Ten Metre Walk (EFIT)

This test was completed at both baseline and session 2 by 20 patients in the frontal lobe tumour group, 9 patients in the temporal lobe tumour group, 2 in the limbic group, 7 in the parietal lobe tumour group, 2 in the occipital lobe tumour group, 11 in the multiple lobes tumour group and 5 patients in the 'other' location group.

The frontal lobe tumour group scored a mean of 7.0 (SD 1.4) at baseline and 7.2 (SD 1.6) at session 2. The temporal lobe tumour group had a mean baseline score of 6.2 (SD 1.2) and a mean session 2 score of 6.5 (SD 1.2). The limbic tumour group has mean scores of 7.4 (SD 2.0) and 6.7 (SD 1.1) at baseline and session 2, respectively. The baseline mean score for the occipital lobe tumour group was 6.1 (SD 0.7) and at session 2 was 7.1 (SD 1.4). The multiple lobes tumour group had a baseline mean score of 6.4 (SD 1.2) and a session 2 mean score of 7.6 (SD 1.8). The 'other' tumour location group scored means of 6.5 (SD 2.2) and 6.5 (SD 1.5) at baseline and session 2, respectively. The mean (standard error) scores at baseline and session 2 for each tumour lobe group are shown in Figure 9.14.

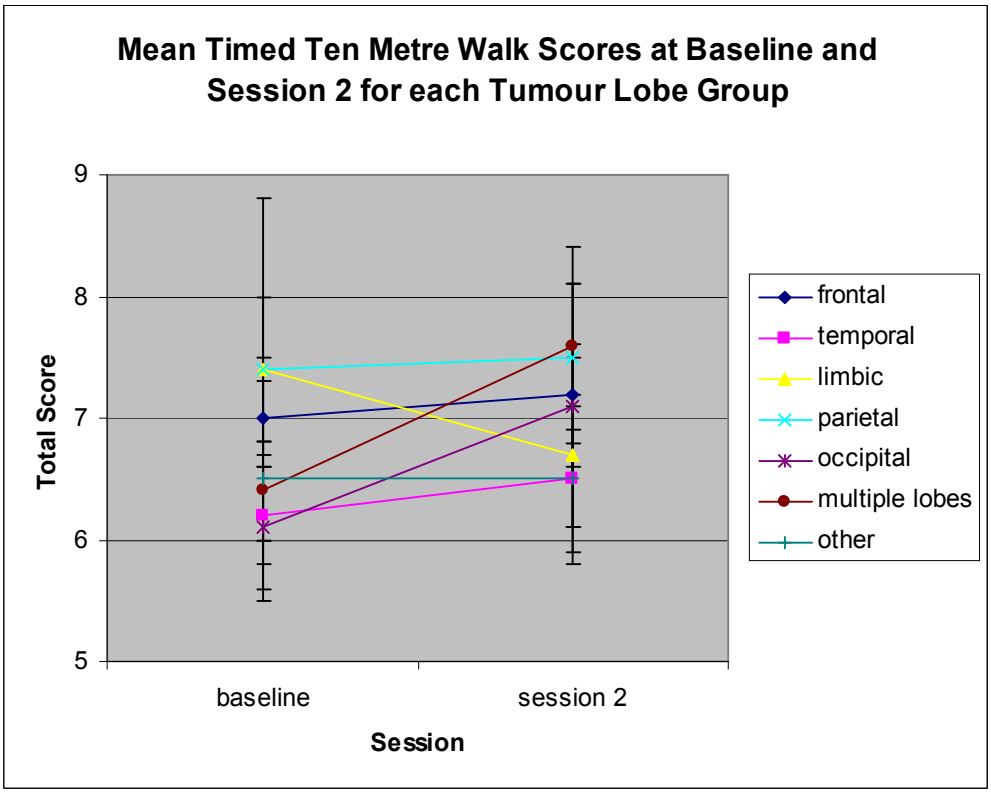


Figure 9.14. Baseline and session 2 timed ten metre walk scores for each tumour lobe group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of the timed ten metre walk test score at baseline, $F(1,41) = 42.022$, $p < 0.001$, partial $\eta^2 = 0.506$. The covariate age had no significant main effect in the model, $F(1,41) = 3.204$, $p = 0.081$, partial $\eta^2 = 0.072$. The covariate NART also had no significant main effect, $F(1,41) = 0.569$, $p = 0.455$, partial $\eta^2 = 0.014$. Sex had no significant effect in the model, $F(1,41) = 0.190$, $p = 0.665$, partial $\eta^2 = 0.005$. Tumour lobe had no significant main effect in the model that included the effects of the covariates age, NART score and baseline timed ten metre walk score, $F(6,41) = 1.318$, $p = 0.271$, partial $\eta^2 = 0.162$. There was a significant interaction between tumour lobe and sex, $F(4,41) = 2.680$, $p = 0.045$, partial $\eta^2 = 0.207$. Female participants in the multiple lobes and 'other' location group took longer to complete the timed ten metre walk than the male participants.

The estimated marginal mean scores for each tumour lobe group, adjusted for age, NART score and baseline timed ten metre walk score – derived from the above described general linear model – were 7.0 (SE 0.2) for the frontal lobe tumour group; 7.0 (SE 0.4) for the temporal lobe tumour group; 6.2 (SE 0.8) for the limbic tumour group; 6.8 (SE 0.4) for the parietal lobe tumour group; 7.6 (SE 0.7) for the occipital lobe tumour group; 7.8 (SE 0.3) for the multiple lobes tumours group and 6.7 (SE 0.5) for the 'other' location tumour group.

Table 9.2. Overview of comparisons of post-operative test performance in each tumour lobe group.

Test	Effect of tumour lobe		
	F	Sig. (p)	n^{2*}
Inspection Time (All Data)	0.495	0.808	0.063
Inspection Time (Valid Data)	0.759	0.585	0.091
Digit Symbol Coding (Total)	0.792	0.581	0.097
EFIT Williams Delayed Recall Test (total)	2.091	0.073	0.214
EFIT Nine Hole Peg Test (Right Hand)	0.641	0.697	0.077
EFIT Nine Hole Peg Test (Left Hand)	0.919	0.491	0.109
EFIT Timed Ten Metre Walk (secs)	1.318	0.271	0.162

n^2 = the proportion of variance accounted for by the covariate (tumour lobe).

9.4 Hemispheric Lateralisation

Thirty-one patients with left hemisphere tumours and 31 patients with right hemisphere tumours were tested at both baseline and session 2, post-operatively.

9.4.1 Inspection Time Scores: All Inspection Time Data

Data from 29 patients with left hemisphere tumours and 29 with right hemisphere tumours was available for analysis.

The left hemisphere tumour group had a mean score of 122.3 (SD 14.4) at baseline and 117.3 (SD 20.6) at session 2. The right hemisphere tumour group scored a baseline mean of 112.2 (SD 25.0) and a mean score at session 2 of 109.5 (SD 24.1). The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.15.

Figure 9.15. Baseline and session 2 inspection time scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of baseline inspection time score in the model, $F(1,51) = 21.401$, $p < 0.001$, partial $\eta^2 = 0.296$. Participants who scored higher on inspection time at baseline were significantly more likely to also score better at session 2. The effect of the covariate age approached significance in the model, $F(1,51) = 3.730$, $p = 0.059$, partial $\eta^2 = 0.068$. The covariate NART score had no significant main effect in the model, $F(1,51) = 1.192$, $p = 0.280$, partial $\eta^2 = 0.023$. Sex had no significant main effect in the model, $F(1,51) = 0.600$, $p = 0.442$, partial $\eta^2 = 0.012$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates age, NART score and baseline inspection time score, $F(1,51) = 0.197$, $p = 0.659$, partial $\eta^2 = 0.004$. There was no significant interaction between tumour hemisphere and sex, $F(1,51) = 0.008$, $p = 0.928$, partial $\eta^2 < 0.001$. Thus, in the presence of the effect of the covariates, there was no significant difference

in the extent of post-operative deterioration between the left and right hemisphere tumour groups.

9.4.2 Inspection Time Scores: Valid Inspection Time Data

Twenty-seven patients with left hemisphere tumours and 23 with right hemisphere tumours had valid baseline inspection time scores and also completed session 2 post-operative testing.

The mean baseline score for the left hemisphere group was 124.4 (SD 11.7) and at session 2 was 117.7 (SD 20.3). The right hemisphere group had a baseline mean score of 122.6 (SD 14.2) and a mean score of 115.7 (SD 19.7) at session 2. The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.16.

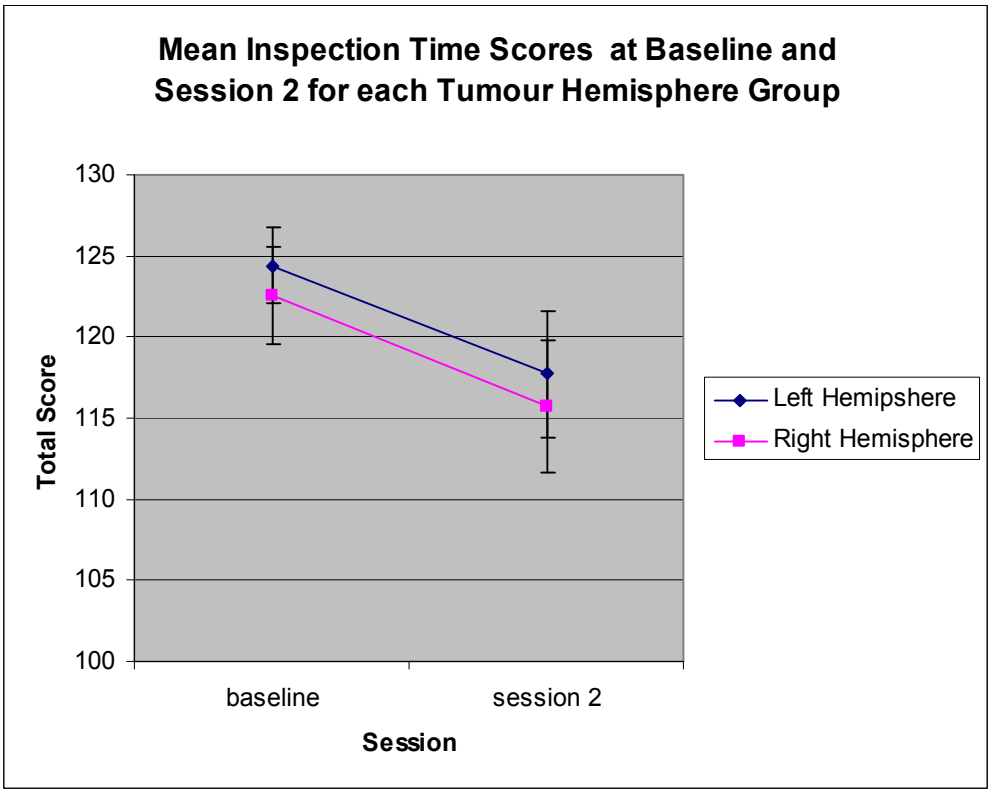


Figure 9.16. Baseline and session 2 inspection time scores (valid baseline data) for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

General linear modelling again revealed a significant main effect of the baseline inspection time score, $F(1,43) = 11.077$, $p = 0.002$, partial $\eta^2 = 0.205$. Participants with higher inspection time scores at baseline were significantly more likely to have higher scores at session 2. There was no significant main effect of the covariates age, $F(1,43) = 2.229$, $p = 0.143$, partial $\eta^2 = 0.049$; or NART score, $F(1,43) = 2.406$, $p = 0.128$, partial $\eta^2 = 0.053$. Sex had no significant main effect in the model, $F(1,43) = 0.035$, $p = 0.851$, partial $\eta^2 = 0.001$. Tumour hemisphere had no significant effect in the model that included the effects of the covariates age, NART score and baseline inspection time score, $F(1,43) = 0.445$, $p = 0.508$, partial $\eta^2 = 0.010$. The interaction between tumour hemisphere and sex was not significant, $F(1,43) = 1.345$, $p = 0.253$, partial $\eta^2 = 0.030$. Thus, in the presence of the effects of the covariates, there was no significant difference

between the left and right hemisphere groups in terms of post-operative inspection time performance when only valid baseline scores were included in the analysis.

9.4.3 Digit Symbol Coding

There were 29 patients in the left hemisphere tumour group and 29 in the right hemisphere tumour group who completed the digit symbol coding score at both baseline and session 2, post-operatively.

The mean baseline score for the left hemisphere tumour group was 61.4 (SD 20.3) and at session 2 was 57.2 (SD 19.6). The right hemisphere tumour group had a mean score of 60.7 (SD 21.3) at baseline and 56.6 (SD 21.7) at session 2. The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.17.

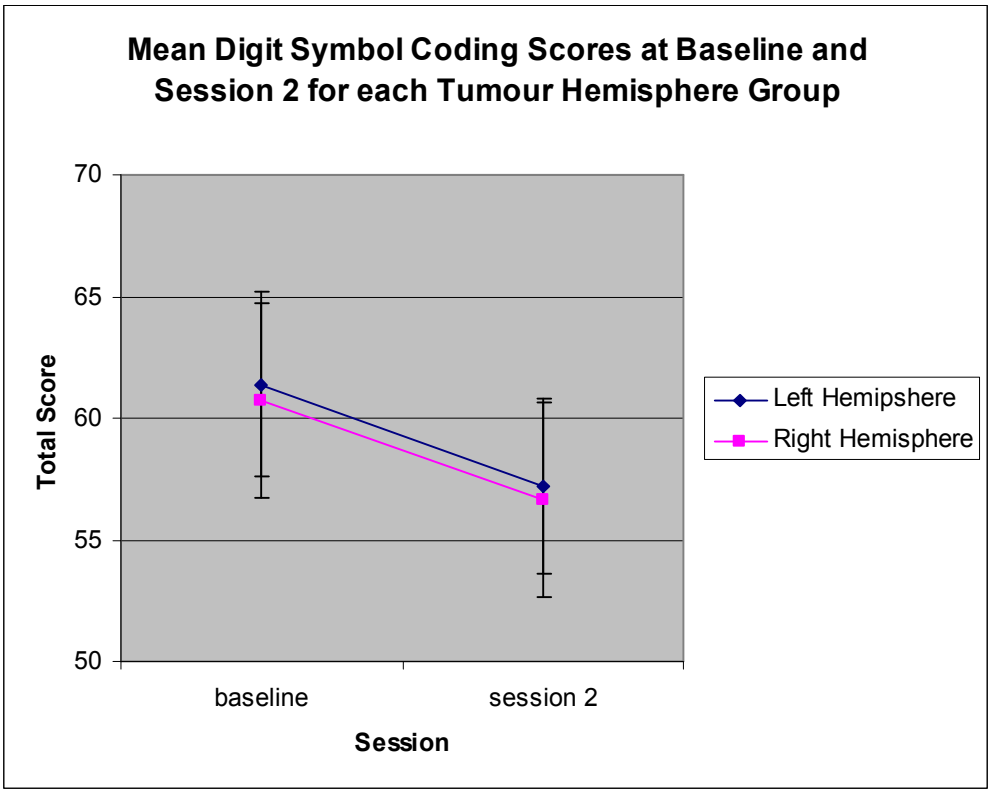


Figure 9.17. Baseline and session 2 digit symbol coding scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of the baseline digit symbol coding score, $F(1,51) = 114.215$, $p < 0.001$, partial $\eta^2 = 0.691$. Patients with higher scores at baseline were significantly more likely to have higher scores when tested at session 2. There was no significant effect of either of the covariates age, $F(1,51) = 0.016$, $p = 0.899$, partial $\eta^2 < 0.001$; or NART score, $F(1,51) = 0.245$, $p = 0.623$, partial $\eta^2 = 0.005$. Sex had no significant main effect in the model, $F(1,51) = 2.851$, $p = 0.097$, partial $\eta^2 = 0.053$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates age, NART score and baseline test score, $F(1,51) = 0.054$, $p = 0.817$, partial $\eta^2 = 0.001$. The interaction between sex and tumour hemisphere was not significant, $F(1,51) = 0.049$, $p = 0.826$, partial $\eta^2 = 0.001$. Thus, in the presence of the effects of age, NART score and baseline score, there was no significant difference between the left and right hemisphere groups in terms of performance on digit symbol coding at session 2.

9.4.4 Williams Delayed Recall Test (EFIT)

Twenty-nine patients with left hemisphere tumours and 31 with right hemisphere tumours completed the Williams delayed recall test at both baseline and post-operatively at session 2.

The mean score at baseline for the left hemisphere tumour group was 8.5 (SD 4.9) and at session 2 was 11.3 (SD 6.8). The right hemisphere tumour group had a baseline mean score of 9.0 (SD 5.8) and a session 2 mean score of 11.2 (SD 5.6). The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.18.

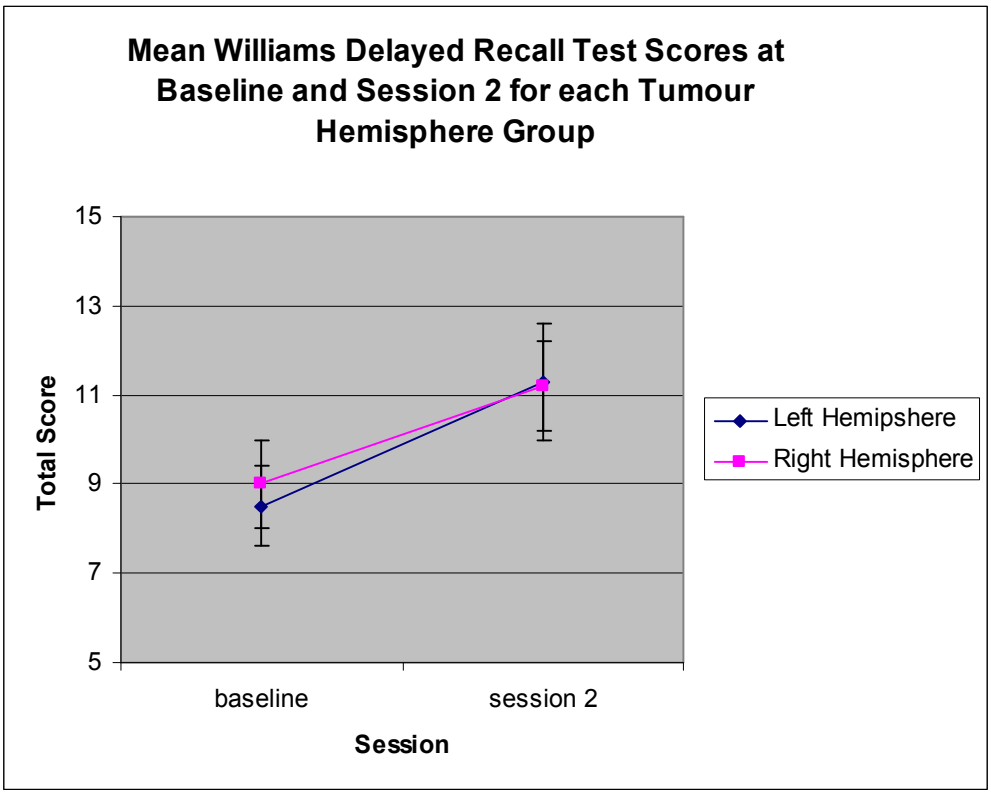


Figure 9.18. Baseline and session 2 Williams delayed recall test scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of the baseline tests score in the model, $F(1,53) = 10.766$, $p = 0.002$, partial $\eta^2 = 0.169$. There was no significant main effect of either of the covariates age, $F(1,53) = 1.213$, $p = 0.276$, partial $\eta^2 = 0.022$; or NART score, $F(1,53) = 0.712$, $p = 0.403$, partial $\eta^2 = 0.013$. Sex had no significant effect in the model, $F(1,53) = 0.007$, $p = 0.931$, partial $\eta^2 < 0.001$. Tumour hemisphere had no significant main effect in the model that included the effects of the covariates age, NART score and baseline Williams delayed recall test score, $F(1,53) = 0.017$, $p = 0.897$, partial $\eta^2 < 0.001$. The interaction between sex and tumour hemisphere was not significant, $F(1,53) = 0.081$, $p = 0.777$, partial $\eta^2 = 0.002$. Thus, in the presence of the effects of the covariates, there was no significant difference between the performance of the left and right hemisphere tumour groups at session 2 post-operatively.

9.4.5 Nine Hole Peg Test (Right Hand, EFIT)

Twenty-nine left hemisphere tumour patients and 31 right hemisphere tumour patients completed this test at both baseline and post-operatively at session 2.

The mean score for the left hemisphere tumour group at baseline was 13.8 (SD 2.6) and at session 2 was 14.3 (SD 1.9). The left hemisphere tumour group had a baseline mean score of 14.0 (SD 3.3) and at session 2 had a mean score of 14.9 (SD 3.5). The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.19.

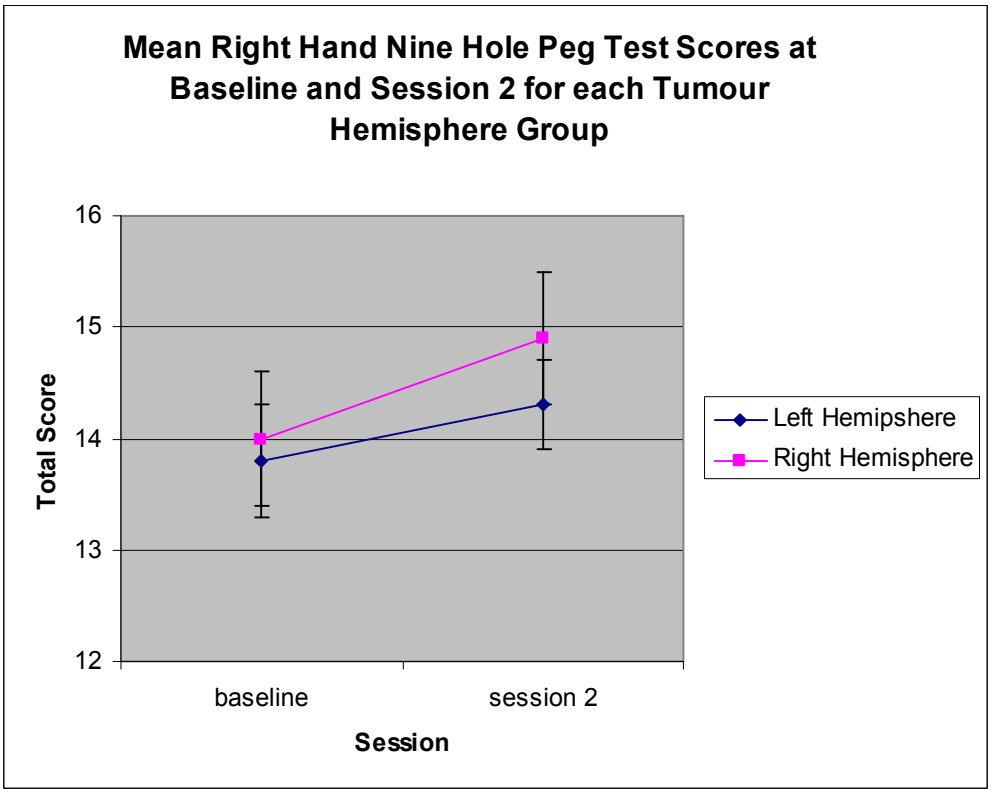


Figure 9.19. Baseline and session 2 right hand nine hole peg test scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

General linear modelling revealed a significant main effect of the baseline right hand nine hole peg test score, $F(1,53) = 41.854$, $p < 0.001$, partial $\eta^2 = 0.441$. Participants with faster scores at baseline were significantly more likely to also have faster scores at session 2. There was no significant main effect of either of the covariates age, $F(1,53) = 1.005$, $p = 0.321$, partial $\eta^2 = 0.019$; or NART score, $F(1,53) = 0.035$, $p = 0.853$, partial $\eta^2 = 0.001$. Sex had no significant main effect in the model, $F(1,53) = 0.089$, $p = 0.766$, partial $\eta^2 = 0.002$. The effect of hemispheric location of the tumour had no significant main effect in the model that included the effects of the covariates age, NART score and baseline score, $F(1,53) = 0.666$, $p = 0.418$, partial $\eta^2 = 0.012$. The interaction between sex and hemisphere was also not significant, $F(1,53) = 1.342$, $p = 0.252$, partial $\eta^2 = 0.025$. Thus, in the presence of the effect of the aforementioned covariates, there was no significant difference between the extent of post-operative deterioration on this test in the left and right hemisphere tumour groups.

9.4.6 Nine Hole Peg Test (Left Hand, EFIT)

Thirty patients with left hemisphere tumours and 29 with right hemisphere tumours completed the left hand nine hole peg test at both baseline and session 2.

The mean baseline score for the left hemisphere tumour group on this test was 14.6 (SD 2.6) and at session 2, post-operatively, was 14.9 (SD 2.7). The right hemisphere tumour group scored a mean of 16.3 (SD 4.2) at baseline and 17.0 (SD 3.8) at session 2. The mean (standard error) scores at baseline and session2 for each hemisphere group are shown in Figure 10.20.

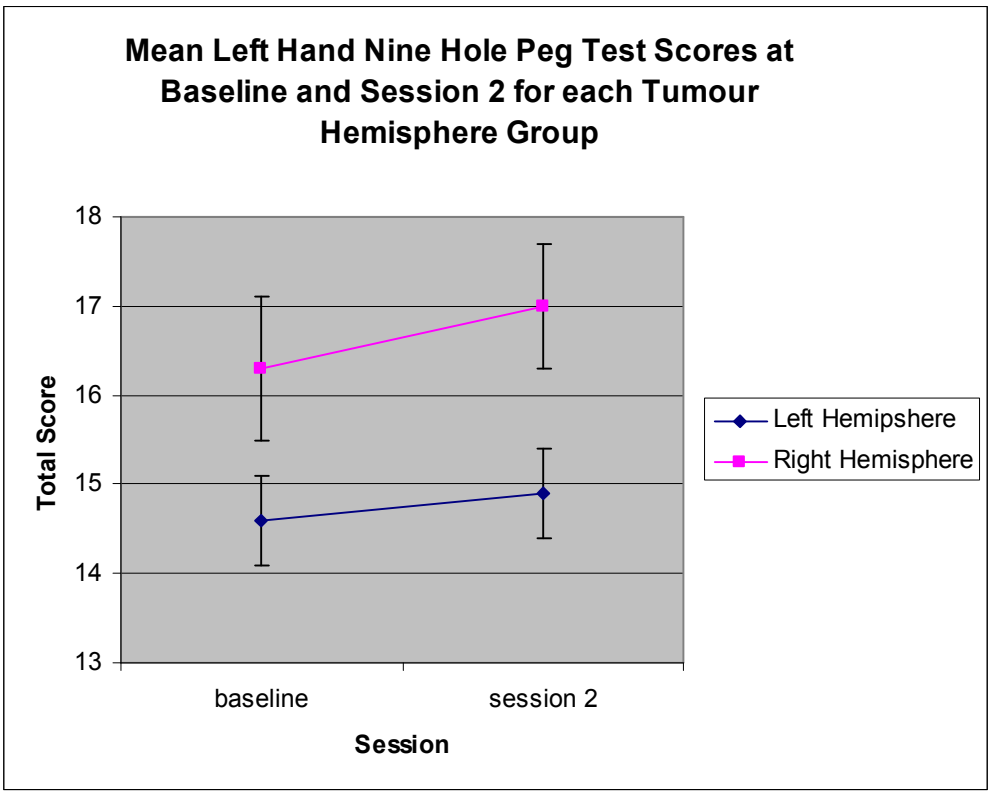


Figure 9.20. Baseline and session 2 left hand nine hole peg test scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

There was a significant main effect of the baseline test score in the general linear model, $F(1,52) = 99.183$, $p < 0.001$, partial $\eta^2 = 0.656$. Those participants who had faster scores at baseline were significantly more likely to have faster scores at session 2. The covariate age also had a significant main effect in the model, $F(1,52) = 7.215$, $p = 0.010$, partial $\eta^2 = 0.122$. The covariate NART score had no significant main effect in the model, $F(1,52) = 1.228$, $p = 0.273$, partial $\eta^2 = 0.023$. The effect of sex was not significant, $F(1,52) = 0.778$, $p = 0.382$, partial $\eta^2 = 0.015$. The effect of hemispheric location of tumour did not reach significance in the model that included the effects of the covariates age, NART score and baseline left hand nine hole peg test score, $F(1,52) = 3.377$, $p = 0.072$, partial $\eta^2 = 0.061$. The interaction between sex and hemisphere was not significant, $F(1,52) = 0.015$, $p = 0.904$, partial $\eta^2 < 0.001$. In the presence of the effects of the covariates, there was no significant difference in the extent of post-operative deterioration on this measure in the left and right hemisphere tumour groups.

9.4.7 Timed Ten Metre Walk (EFIT)

There were 27 patients with left hemisphere tumours and 27 patients with right hemisphere tumours who completed the timed ten metre walk at both baseline and session 2, post-operatively.

The mean baseline score for the left hemisphere tumour group was 6.7 (SD 1.5) and the mean score at session 2 was 6.9 (SD 1.7). The right hemisphere tumour group had a baseline mean score of 6.9 (SD 1.3) and a session 2 mean score of 7.4 (SD 1.7). The mean (standard error) scores at baseline and session 2 for each hemisphere group are shown in Figure 9.21.

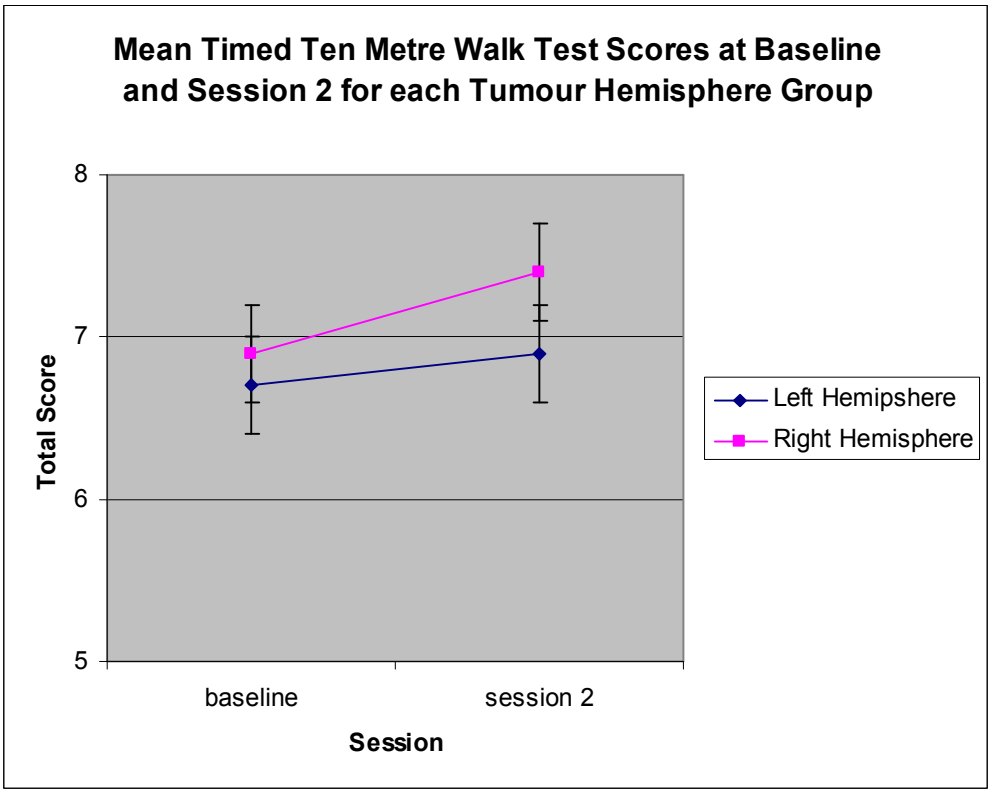


Figure 9.21. Baseline and session 2 timed ten metre walk test scores for each tumour hemisphere group. Points are raw mean scores. Error bars: standard error means.

There was a significant main effect of the timed ten metre walk test score at baseline in the general linear model, $F(1,47) = 28.629$, $p < 0.001$, partial $\eta^2 = 0.379$. Those participants with faster scores on the test at baseline were significantly more likely to have faster scores at session 2 testing. The effect of the covariate age was also significant in the model, $F(1,47) = 5.548$, $p = 0.023$, partial $\eta^2 = 0.106$. Older participants were more likely to take longer to complete the test. The covariate NART score had no significant min effect in the model, $F(1,47) = 0.668$, $p = 0.418$, partial $\eta^2 = 0.014$. Sex had no significant effect in the model, $F(1,47) = 0.113$, $p = 0.738$, partial $\eta^2 = 0.002$. The effect of tumour hemisphere was not significant in the model that included the effects of the covariates age, NART score and baseline test score, $F(1,47) = 1.162$, $p = 0.287$, partial $\eta^2 = 0.024$. The interaction between sex and tumour hemisphere was not significant, $F(1,47) = 1.519$, $p = 0.224$, partial $\eta^2 = 0.031$. Thus, in the presence of the

effects of the covariates, there was no significant difference between the post-operative performance of the left and right hemisphere tumour groups.

Table 9.3. Overview of comparisons of post-operative test performance in the left and right hemisphere groups.

Test	Effect of tumour hemisphere		
	F	Sig. (p)	n^{2}*
Inspection Time (All Data)	0.197	0.659	0.004
Inspection Time (Valid Data)	0.445	0.508	0.010
Digit Symbol Coding (Total)	0.054	0.817	0.001
EFIT Williams Delayed Recall Test (total)	0.017	0.897	< 0.001
EFIT Nine Hole Peg Test (Right Hand, secs)	0.666	0.418	0.012
EFIT Nine Hole Peg Test (Left Hand, secs)	3.377	0.072	0.061
EFIT Timed Ten Metre Walk (secs)	1.162	0.287	0.024

* n^{2} = the proportion of variance accounted for by the tumour hemisphere.

9.5 Session 3 Follow-Up: Low-Grade Glioma vs. High-Grade Glioma

9.5.1 Rationale and Overview of Analysis Procedure

Only patients with either a low-grade or high-grade glioma participated in session 3 follow-up testing and many patients completed only inspection time testing at this time due to time constrains (see Chapter 3.4). Therefore, a preliminary analysis was carried out to examine whether or not the deterioration in inspection time performance in the immediate post-operative period was transient, and whether the performance of the low-grade and high-grade glioma groups differed. The following analysis was carried out by Mike Allerhand, Statistician in the Department of Psychology, University of Edinburgh.

9.5.2 Results

The mean (standard deviation) inspection time score for the low-grade glioma group was 126.1 (SD 11.5) at baseline, 118.2 (SD 23.4) at session 2 and was 129.1 (SD 9.2) at the time of session 3 follow-up. The mean inspection time score for the high-grade glioma group was 110.0 (SD 21.8) at baseline, 113.3 (SD 22.5) at session 2 testing, and 116.0 (SD 20.5) at session 3.

To compare the performance of the low-grade and high-grade glioma cohorts across the three testing sessions, the difference in trajectories of the mean inspection time scores over the three measurement occasions between the two groups (low grade and high grade glioma) was tested using a quadratic multi-level growth model (using the lme4 package in R), since this allowed inclusion of inspection time scores from all participants, not only scores from those patients who had been tested across all three sessions:

$$
\begin{aligned}
y = &\quad \pi_0 + \pi_1 TIME + \pi_2 TIME^2 + e \\
\pi_0 = &\quad \gamma_{00} + \gamma_{01} TUMOURTYPE + \zeta_0 \\
\pi_1 = &\quad \gamma_{10} + \gamma_{11} TUMOURTYPE + \zeta_1 \\
\pi_2 = &\quad \gamma_{20} + \gamma_{21} TUMOURTYPE + \zeta_2
\end{aligned}
$$

The level-2 model is conditioned on a dichotomy TUMOURTYPE, (low and high-grade). The parameter of interest is the fixed effect γ_{21} that describes the difference in the second-order growth parameter (effectively the curvature of the growth curve) between the two groups (low and high). The estimate is significant (p=.024). Figure 9.22 shows that the low-grade glioma group has a curved trajectory, with session 2 scores lower than both baseline and session 3, whereas the high-grade glioma group has a straight trajectory. The significant p-value suggests that the difference between the performance of the two groups is different: the low-grade glioma group deteriorated in the initial post-operative period, but maintained their baseline performance when tested at session 3, whereas the high-grade glioma group did not show this pattern of transient deterioration. The implications of this are discussed in the following section.

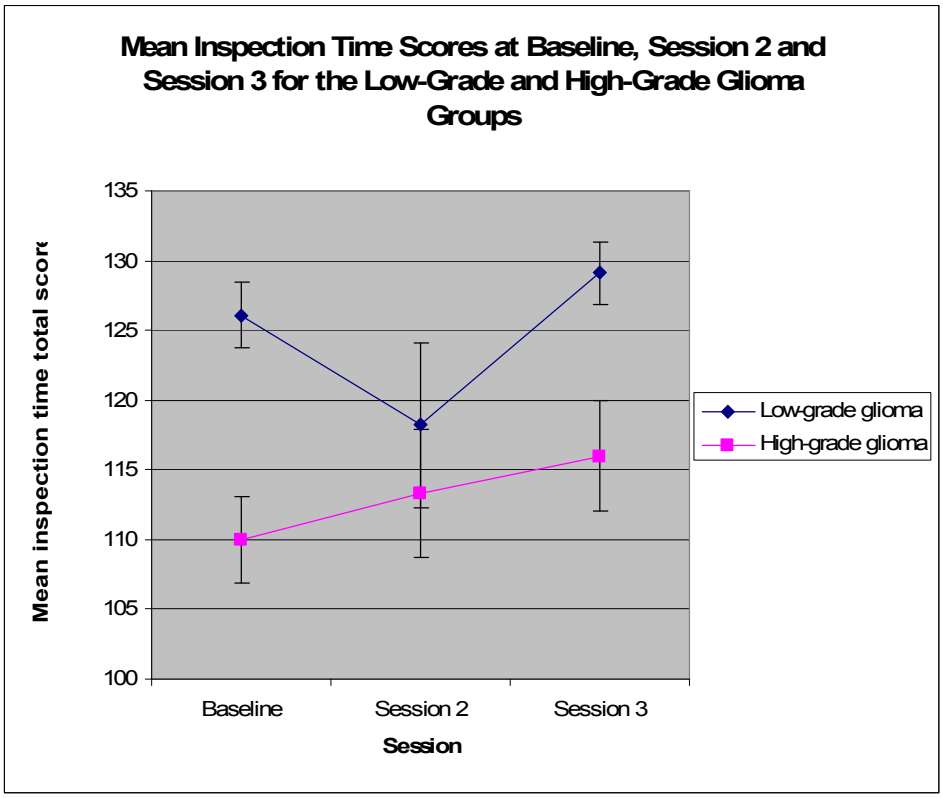


Figure 9.22. Baseline, session 2 and session 3 inspection time scores for the low-grade and high-grade glioma groups. Points are raw mean scores. Error bars: standard error means.

9.6 Tumour Characteristics: Discussion

9.6.1 Tumour Histology

Overall, there was no significant main effect of tumour type on any of the baseline (pre-operative) test scores. However, post-hoc analyses revealed that, when all inspection time data was included in analysis, the performance of the high-grade glioma group was significantly worse than that of the low-grade glioma group, although this difference was eliminated when invalid inspection time scores were excluded. No other significant differences between the performance of the low-grade and high-grade glioma groups were found.

Following surgery, at the initial post-operative testing session, there was no main effect of tumour type on follow-up test performance. However, post-hoc tests revealed that the low-grade glioma group deteriorated significantly more than the high-grade glioma group on the Williams Delayed Recall Test. On the inspection time measure, there was a trend towards greater deterioration in the low-grade glioma group following surgery when all data was included in the analysis and also when only valid data was included. However, these differences were not statistically significant. Preliminary analysis of the performance of the low-grade and high-grade glioma tumour groups on the inspection time task across all three testing sessions suggested that the initial post-operative deterioration observed in the low-grade glioma group may be transient, with the low-grade glioma group exceeding baseline inspection time scores when tested on a third occasion.

That there were so few differences between the baseline performance of the tumour type groups on any of the cognitive measures is somewhat surprising. It has generally been reported that high-grade glioma patients tend to present with more severe neurological impairment than their low-grade counterparts, given the more aggressive nature of their tumour (DeAngelis, 2001). Thus, it was expected that the high-grade glioma group would exhibit more severe cognitive impairment in the pre-operative period than the low-grade glioma group. However, since the majority of

studies examining cognition in brain tumour patients tend to focus on a single, specific histological group, this prevents direct comparison of the extent of impairment in different histological groups as measured by the same tests, at the same point in the disease journey. In one of the few studies that directly compare different tumour types, patients with rapidly growing tumours (i.e. high-grade gliomas) were found to exhibit more severe cognitive impairment than patients with more slowly-growing tumours (Hom and Reitan, 1984). Conversely, in a comparison of a group of patients with an anaplastic astrocytoma with a group of patients with a glioblastoma multiforme, no significant differences between the two groups on a battery of tests assessing five domains of cognition were found (Kayl and Meyers, 2003). This finding suggests that patients with more malignant disease do not suffer more severe cognitive deficits than those with less aggressive tumours, although the study did not a directly compare high and low-grade glioma patients. In the present study, the high-grade glioma group did perform significantly less well than the low-grade glioma group on inspection time at baseline, but only when all data was included in analyses. That this difference was eliminated when invalid scores were removed from the data set suggests that high-grade glioma patients were significantly more likely to have invalid scores than the low-grade glioma group. The relatively high proportion of high-grade glioma patients who failed to achieve validity criterion on the inspection time test may not only reflect poorer concentration and comprehension of the task demands in the high-grade glioma group but could also reflect a slowing of visual information processing in this group to such an extent that validity criterion could not be met despite full comprehension of the task. Thus, the difference between the two glioma groups when all inspection time data is analysed may accurately reflect a greater extent of visual information processing slowing in high-grade glioma patients in particular. It is for this reason that an alternative version of the inspection time task using an adaptive staircase procedure (as discussed in Chapter 5.7) may prove particularly useful since this would allow us to determine whether visual information processing is slowed to such an extent in high-grade glioma patients that validity criterion on the more commonly used version of the task cannot be met in some patients with highly aggressive tumours. Moreover, the poorer performance of the high-grade glioma group on verbal fluency (by

comparison with the low-grade glioma and metastasis groups) and on the Rey Auditory Verbal Learning Test (by comparison with the meningioma and 'other' tumour groups) suggests that the extent of cognitive impairment was greater in the patients with highly aggressive tumours. As highlighted in Chapter 4, a significantly lower proportion of high-grade glioma patients than expected were entered into the study and, conversely, the majority of low-grade glioma patients did take part. This selection bias must be taken into account when interpreting the results comparing the performance of the different tumour type groups at baseline. Since those patients who had the most severe impairment and were unable to consent or unwilling to take part in the study were more likely to have a high-grade glioma, the cognitive impairment in the high-grade glioma group as a whole is likely to be greater than that observed in the present study, as a result of inclusion bias.

Comparatively few studies have examined cognition in patients with meningiomas and brain metastases and the present study provides further insight into this under-studied area. There was no consistent pattern of either greater impairment or superior performance in either of these tumour type groups on the baseline cognitive test battery by comparison with the other tumour type groups. In one of the few studies of cognition in patients with brain metastases, a specific impairment on tests of recall and delayed memory was found (Herman et al., 2003). However, in the present study, the metastasis group were no more impaired on the baseline measures of memory and recall (e.g. Rey Auditory Verbal Learning Test and Williams Delayed Recall Test) than the other tumour type groups. In contrast with glial tumours, meningiomas tend to be benign, non-infiltrating tumours and it was expected that patients with a meningioma would therefore exhibit less severe impairment than those patients with gliomas, particularly those of a high-grade. The only baseline test on which the meningioma group significantly outperformed any other tumour group was the Rey Auditory Verbal Learning Test, yet these patients showed a similar degree of impairment as the other tumour type groups on the other baseline tests. In one of the few studies to specifically examine cognition in meningioma patients prior to surgery, Tucha et al. (2003) report significant impairment in several cognitive domains in a group of patients assessed pre-operatively by comparison with a

matched group of healthy control participants. Thus, the finding that meningioma patients showed similar levels of impairment as the other tumour type groups in the present study lends support to Tucha et al.'s (2003) conclusions that suggest cognitive impairment is common to meningioma patients prior to surgical intervention.

There were no differences between any of the tumour type groups with respect to the levels of anxiety and depression as measured by the Hospital Anxiety and Depression Scale (HADS). Thus, it can be reasonably assumed that the potential for emotional distress to negatively impact cognitive test performance was similar in each tumour type group. However, this finding is inconsistent with a study that found patients with a meningioma had higher levels of both anxiety and depression pre-operatively, as measured by the HADS (Pringle et al., 1999).

At the initial post-operative testing session (session 2), there was no significant overall effect of tumour type on patient performance on any of the follow-up measures. This suggests that there is no single group of patients with a specific tumour type who have an increased risk of suffering greater cognitive impairment following surgical intervention. However, although there was no statistically significant effect, the low-grade glioma group did show a trend towards greater deterioration following surgery by comparison with the high-grade glioma group on a number of measures, including inspection time, digit symbol coding and the Williams Delayed Recall Test. Although studies have examined post-operative cognitive impairment in low-grade glioma patients specifically, no studies have assessed the risk of impairment by comparison with other tumour type groups. The present results suggest that there may be a tendency for low-grade glioma patients to deteriorate to a greater extent than high-grade glioma patients following surgery The lack of significant main effect in these analyses may be a result of reduced power to detect differences due to relatively high levels of post-operative attrition in the brain tumour group that makes type 2 statistical error more likely. Although the low-grade glioma group appeared to be at greater risk of deterioration at the time of the initial post-operative testing, that the low-grade group exceeded their baseline scores on

inspection time testing when assessed for a third time, 10-14 days later, suggests that any post-operative deterioration may be transient. Future studies could specifically compare the relative risk of impairment in low and high-grade glioma patients in a larger cohort to overcome the problems associated with high levels of attrition in the present study. Previous studies examining post-operative neurological and cognitive complications in low-grade glioma patients have found evidence to suggest that, three months after surgery, patients may actually exceed their pre-operative function (Duffau et al., 2003, Teixidor et al., 2006) and this is supported by the findings of the present study. However, further investigation is necessary, with a larger battery of tests administered at session 3 given that it was only possible to analyse inspection time scores at session 3 due to failure of the majority of patients to complete the remaining follow up measures. Moreover, those glioma patients who were assessed at session 3 are likely to have been those patients who had a more positive outcome following surgery.

The finding that the group of patients with a meningioma did not deteriorate to a lesser extent post-operatively than the other tumour type groups contradicts the results of the aforementioned study (Tucha et al., 2003). These authors found that the cohort of meningioma patients actually improved post-operatively on some tests assessing several aspects of cognition, and maintained their pre-operative performance on many other measures. However, the cohort of meningioma patients recruited into the present study had tumours located throughout the brain, as opposed to only frontal lobe tumours. Additionally, there was no adequate control group recruited into the study by Tucha et al., (2003) which questions the validity of their findings. Thus, since no direct comparison with other tumour type groups was carried out and the test battery comprised measures that were perhaps less suitable for repeated assessment than those used in the present study, the findings reported in this thesis question the conclusions made by Tucha et al. (2003).

Future studies could seek to confirm the present findings in a larger cohort of patients with meningiomas. Moreover, since the group of patients with ‘other’ tumours was relatively small and consisted primarily of patients with pituitary

adenomas, it would be of interest to specifically examine cognition in patients with pituitary tumours. Little research had been directed towards this area and given that, in this very small sample of patients with these relatively benign pituitary tumours, cognition appeared to be disrupted in both the pre- and post-operative periods, future research could seek to confirm this finding in a much larger sample of patients.

9.6.2 Hemispheric Location

Very few significant differences between the left and right hemisphere tumour groups on the baseline measures were found. When all inspection time data was analysed, the left hemisphere group significantly outperformed the right hemisphere group, however this difference was eliminated when invalid scores were removed from the model. The left hemisphere group were significantly impaired by comparison with the right hemisphere tumour group on the RAVLT and the right hemisphere group were impaired on the left hand nine hole peg test compared with the left hemisphere group. There were no other significant differences between the left and right hemisphere tumour groups on any of the other baseline tests and the two groups did not differ at all in terms of the extent of post-operative decline. The number of patients in each hemisphere group was evenly distributed, and any patients with bi-hemispheric tumours, or tumours located elsewhere (e.g. on the pituitary gland) were excluded from the analyses in this instance.

The literature to date has been somewhat conflicting regarding the effect of hemispheric location on cognitive function in patients with brain tumours. Studies have generally reported that right hemisphere lesions are associated with impairment on tests that rely on visual-perceptual skills (Scheibel et al., 1996, Klein et al., 2001), whereas left hemisphere tumours tend to result in poorer performance on verbal learning and language tasks (Hom and Reitan, 1984, Scheibel et al., 1996). Thus, based upon this evidence, it was expected that the left hemisphere group would perform significantly less well on verbal fluency and the Rey Auditory Verbal Learning Test (RAVLT) since both measures involve language function and the RAVLT is a measure of verbal learning. However, relative impairment in the left

hemisphere group was observed on the RAVLT only, with no significant difference found between the two hemisphere groups on the verbal fluency measure. Patients with tumours in the right hemisphere were significantly more likely to have invalid inspection time scores, given that the group were significantly impaired by comparison with the left hemisphere group when all inspection time data was included in the analysis. This may not simply reflect a lack of comprehension of task instructions but could again be the result of impaired visual perceptual skills that precluded successful completion of even the longest stimulus presentation trials. However, studies of the functional anatomy of the inspection time task have found that both left and right hemispheric brain regions are involved in carrying out the task (Deary et al., 2001). Therefore, given that both sides of the brain are involved in inspection time task performance, this cannot explain the poorer performance of the right hemisphere group on the test.

That there was no significant difference between the left and right hemisphere groups in terms of the extent of impairment on the Williams Delayed Recall Test contrasts with the findings of an initial investigation into the utility of the EFIT tests in neuro-oncological patients. Grant et al. (1994) found that, in a cohort of patients with cerebral gliomas, left hemisphere tumours were associated with poorer memory function on this measure. However, the heterogeneous nature of the brain tumour group in the present study contrasts with the glioma cohort recruited into the aforementioned study. This may have contributed to the lack of differences between the two hemisphere groups, since the brain tumour group recruited into the present study included patients with benign and less invasive tumours. As expected, the right hemisphere group were significantly slower to complete the nine hole peg test using the left hand at baseline. However, there were no significant differences between the two groups on the right hand version of the test. In their study, Grant et al. (1994) found that the nine hole peg test could detect minor differences between each hand that reflect the lateralisation of the tumour. Thus patients with right hemisphere tumours were significantly slower on the left hand nine hole peg test and vice versa. The present study supports this finding, for the left hand subtest at least. That no

difference on the right hand subtest was found may also be the result of the inclusion of non-glioma patients in the study cohort.

Post-operatively, the extent of decline did not differ between the left and right hemisphere tumour groups on inspection time and other tests, suggesting that there is no increased risk of neither left nor right hemisphere tumour surgery in terms of increasing the extent of cognitive impairment as assessed by the post-operative test battery. There has been little attention directed towards the comparative post-operative cognitive outcome in patients with right vs. left hemisphere tumours. Studies of survival in high-grade glioma patients have reported no significant effect of tumour hemisphere (Jeremic et al., 1994, Kreth et al., 1993) and therefore it is perhaps unsurprising that neither hemisphere group were found to be at greater risk of increased cognitive deterioration following surgery.

9.6.3 Tumour Location: Lobe

There was no consistent pattern of either greater impairment or better performance on the baseline test battery in any of the tumour lobe groups. The occipital lobe group performed significantly less well than each of the other groups on inspection time when all data was included. These differences were only observed when all inspection time scores, including invalid ones, were included in the analysis. When only valid data was examined, the frontal lobe group performed significantly better than the parietal lobe group. There was a significant overall effect of tumour lobe in the model for the RAVLT and Letter-Number Sequencing task, with the group of patients with tumours located elsewhere ('other') performing significantly better on both measures than the majority of the other lobe groups. The parietal lobe group were significantly slower to complete the right hand nine hole peg test than each of the other lobe groups and the occipital lobe group were significantly impaired on the Williams Delayed Recall Test. The effect of tumour lobe approached significance in the model for the digit symbol coding test, with the parietal lobe group performing significantly less well than the majority of the other lobe groups.

Post-operatively, there was no main effect of tumour lobe for any of the test models. However, the temporal lobe group deteriorated to a significantly greater extent than the frontal and parietal lobe groups on the Williams Delayed Recall Test. The group of patients with a tumour invading more than one lobe ('multiple lobes') showed greater relative deterioration than the frontal, parietal and 'other' location groups on the timed ten metre walk. No other significant differences were observed between the different lobe groups.

Few studies have explicitly assessed the comparative cognitive effects of brain tumours dependent upon the lobe in which it is located in both the pre- and post-operative periods. It must be highlighted, however, that the number of patients in the limbic and occipital lobe groups was relatively small, likely as a result of the comparative rarity of tumours located in the limbic system and the increased likelihood of severe visual impairment in patients with occipital tumours that would have rendered these patients ineligible. Therefore, the power to detect any differences dependent upon tumour location is likely to have been reduced. However, these results provide interesting preliminary results upon which future studies could be based.

There were no consistent findings with regard to the effect of tumour lobe on baseline test scores. That the occipital lobe group were significantly worse than all other lobe groups on inspection time, only when all data was included in the analysis, is not unexpected. Given that tumours located within the occipital lobes often result in visual impairment (quadrantanopia or hemianopia), the occipital lobe group were more likely to have difficulty in discriminating between the two inspection time stimulus shapes. As such, these patients were more likely to have invalid inspection time scores. This may also explain the poorer Williams Delayed Recall Test scores in the occipital lobe tumour group since some patients may have had greater difficulty seeing the images to be memorised, resulting in poorer recall. The absence of any other significant differences in the performance of the tumour lobe groups on inspection time testing is likely because relative activation and deactivation in a number of different brain areas have been shown in fMRI studies to be involved

during task performance (Deary et al., 2001, Deary et al., 2004). Therefore, since the neural correlates of inspection time performance are not confined to a single region, this may explain why the different tumour location groups exhibited a similar degree of impairment on the inspection time measure.

Frontal lobe tumours are often associated with a better prognosis, at least in high-grade glioma patients and by comparison with temporal and parietal lobe tumours (Simpson et al., 1993). Since studies have found cognitive function to be a predictor of survival in patients with recurrent gliomas (Meyers et al., 2000), it may be reasonable to expect that those patients with tumours located within the frontal lobes may have outperformed those patients with tumours located elsewhere in the brain at the time of baseline testing. However, several studies that have investigated the neural correlates of different cognitive measures have often reported frontal lobe involvement during task performance. For example, numerous studies have found that the prefrontal cortex is specifically involved in verbal fluency tasks (Parks et al., 1988, Hirshorn and Thompson-Schill, 2006) and studies have found poorer scores on this measure to be associated with frontal lesions, and left frontal lesions in particular may cause greater impairment (Miceli et al., 1981). Increased frontal lobe activity has also been observed during the digit-symbol coding test (Nakahachi et al., 2008), although performance on this measure has also been reported to be negatively affected by the presence of a tumour, regardless of lesion location (Lezak et al., 2004). Studies of trail making test performance in brain-injured populations have found evidence to suggest that frontal lobe activation occurs during task performance (Segalowitz et al., 1992). These findings, coupled with the fact that frontal lobe damage is commonly reported to have a negative impact on executive functioning, would suggest that frontal lobe patients may have been more likely to experience greater impairment on the baseline test battery, compared with patients with tumours located in different lobes throughout the brain. However, the present study does not confirm this theory, given that the frontal lobe tumour group were no more or less impaired on the majority of baseline tests, with the exception of digit symbol coding, on which they outperformed a number of other lobe groups. This therefore questions the proposal that the frontal lobes are specifically involved during performance of the

digit symbol coding test (Nakahachi et al., 2008). Future studies could involve fMRI scanning in brain tumour patients during performance of cognitive tests in order to provide greater insight into the disruptive effect of a tumour on the brain's functional anatomy.

Post-operatively, there were very few significant differences in the extent of deterioration dependent upon the location of the tumour, by lobe. This suggests that the cognitive risks of brain tumour surgery are generally no greater for any specific lobe group. However, the temporal lobe tumour group did exhibit greater deterioration on the Williams Delayed Recall Test when compared with the frontal and parietal lobe groups, suggesting that surgery for temporal lobe tumours may be associated with greater risk of delayed memory impairment than surgery for frontal and parietal lobe tumours, although this may be transient. This finding can likely be explained by the fact that short term memory is related to temporal lobe function, with previous studies reporting that visual recognition memory is mediated by a neural system that included the medial temporal lobe structures and the inferotemporal cortex (Owen et al., 1995).

Thus, overall there were found to be very few significant differences between the different tumour lobe groups on the cognitive measures in neither the pre, nor post-operative period. Future studies involving fMRI scanning in this patient group during performance of the inspection time, and other cognitive measures could provide information regarding the specific brain areas involved in performance of different cognitive tasks. Additionally, further research could also provide greater insight into the extent of disturbance to functional brain networks that occurs as a result of the presence of a tumour that may explain cognitive deficits in neuro-oncological patients that cannot simply be explained as resulting from the location of the tumour (see Chapters 1.4 and 6.3).

10 Quality of Life

10.1 Method of Scoring

The brain tumour and spinal surgery cohorts were administered the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core-30 version 3.0 (EORTC QLQ-C30) at the time of baseline (pre-operative) testing. One hundred and two patients in the brain tumour group and 77 patients in the spinal surgery group completed the questionnaire at baseline. The EORTC QLQ-C30 comprises five functional scales – physical, role, emotional, cognitive and social; three symptom scales - fatigue, nausea and vomiting and pain; six single-item scales – dyspnoea (shortness of breath), insomnia, appetite loss, constipation, diarrhoea and financial impact of disease; and an overall global health scale (health-related quality of life). There are a total of 30 questions in the EORTC QLQ-C30. The questionnaire items were scored according to the procedures recommended by the EORTC Quality of Life Group (Fayers et al., 2001). Briefly, an average score for each scale was calculated (raw score). Each raw score was converted using a linear transformation to a scale that standardises the raw score. Scores ranged from 0 to 100 with higher scores either representing a 'better' level of functioning on the functional and global health scales, or a 'worse' level of symptoms on the symptom and single-item scales.

Both the brain tumour and spinal surgery groups were also administered the Brain Cancer Module (QLQ-BN20), a 20 item questionnaire, intended primarily for brain tumour patients undergoing chemotherapy and/or radiotherapy. It includes four scales assessing future uncertainty, visual disorder, motor dysfunction and communication deficit, and seven single items that measure other disease symptoms (headache, seizures, daytime drowsiness, hair loss, itching, weakness of both legs and bladder control). One hundred and one brain tumour patients and 77 spinal patients completed the QLQ-BN20. Standardised scores were calculated from raw scores, using a linear transformation in the same manner as described for the EORTC

QLQ-C30, with higher scores representing a higher ('worse') level of symptoms (Fayers et al., 2001).

This chapter compares the brain tumour and spinal surgery patient groups on the EORTC QLQ-C30 and QLQ-BN20 in order to determine whether quality of life is significantly worse in patients with brain tumours by comparison with matched surgical controls. Scores on both questionnaire scales are then correlated with inspection time scores at baseline for each group in order to determine whether impairment in specific aspects of quality of life are significantly related to poorer visual information processing.

Given that the scores for several of the EORTC QLQ C30 and QLQ-BN20 were found to be either positively or negatively skewed, non-parametric comparisons were run to compare the brain tumour and spinal surgery groups on each scale.

10.2 Quality of Life: Brain tumour patients vs. spinal surgery controls

The scores for the brain tumour and spinal surgery control group on each scale of the EORTC QLQ-C30 were compared using the Mann-Whitney test. The mean (standard deviation) and median (range) baseline scores for the two groups along with relevant p-values, are shown in Table 10.1 for each scale on the questionnaire. The brain tumour patients did not differ significantly from the spinal surgery patients on the emotional and cognitive function scales, $U = 3496.500$, $z = -1.264$, $p = 0.206$ and $U = 3709.000$, $z = -0.648$, $p = 0.517$, respectively. There was also no significant difference between the levels of nausea and/or vomiting, $U = 3876.500$, $z = -0.173$, $p = 0.862$; dyspnoea, $U = 3887.000$, $z = -0.151$, $p = 0.880$; or diarrhoea, $U = 3590.000$, $z = -1.474$, $p = 0.140$, experienced in the two groups.

However, the brain tumour group had significantly better physical functioning than the spinal surgery patients, $U = 2130.000$, $z = -5.454$, $p < 0.001$. They also had significantly better role (i.e. daily life) function, $U = 2573.000$, $z = -3.993$, $p < 0.001$

and significantly better social function than the spinal surgery patients, U = 2189.000, z = -5.137, p < 0.001.

In terms of symptoms, the spinal surgery group reported significantly more fatigue, U = 2178.000, z = -5.143, p < 0.001; pain, U = 776.500, z = -9.334, p < 0.001; insomnia, U = 3019.000, z = -2.736, p = 0.006; appetite loss, U = 3243.000, z = -2.329, p = 0.020; constipation, U = 3350.500, z = - 2.117, p = 0.034, and financial difficulties, U = 3125.000, z = -2.624, p = 0.009, than the brain tumour cohort. The brain tumour group had significantly better self-rated overall health-related quality of life than the spinal surgery patients, U = 2915.000, z = -2.976, p = 0.003.

Table 10.1. Mean scores and comparisons of the brain tumour and spinal surgery groups on each EORTC QLQ-C30 scale. p-value from Mann-Whitney test. [1] = higher scores indicate better function. [2] = higher scores indicate greater symptom burden.

	Brain Tumour		**Spinal Surgery**		
	Mean (SD)	**Median (range)**	**Mean (SD)**	**Median (range)**	**p value for difference**
Physical Function[1]	80.0 (30.0)	100.0 (0-100)	59.7 (2.8)	60.0 (0 – 100)	< 0.001
Role Function[1]	59.5 (34.0)	66.7 (0 – 100)	38.7 (3.4)	33.3 (0 – 100)	< 0.001
Emotional Function[1]	68.1 (22.3)	75.0 (0 – 100)	64.3 (2.5)	66.7 (0 -100)	0.206
Cognitive Function[1]	69.1 (27.3)	66.7 (0 – 100)	68.0 (2.8)	66.7 (0 – 100)	0.517
Social Function[1]	67.3 (30.8)	66.7 (0 – 100)	43.1 (3.4)	50.0 (0 -100)	< 0.001
Fatigue[2]	32.5 (26.1)	33.3 (0 – 100)	51.8 (2.6)	50.0 (0 – 100)	< 0.001
Nausea/Vomiting[2]	10.8 (17.8)	0.0 (0 – 66.7)	8.9 (1.6)	0.0 (0 – 66.7)	0.862
Pain[2]	20.8 (27.0)	16.7 (0 – 100)	71.6 (2.7)	66.7 (0 – 100)	< 0.001
Dyspnoea[2]	11.1 (21.2)	0.00 (0 – 100)	11.3 (2.3)	0.0 (0 - 66.7)	0.880
Insomnia[2]	40.5 (38.0)	33.3 (0 – 100)	56.3 (4.2)	66.7 (0 – 100)	0.006
Appetite Loss[2]	14.1 (25.4)	0.00 (0 – 100)	21.6 (3.1)	0.0 (0 -100)	0.020

Constipation²	10.1 (20.9)	0.00 (0 – 100)	21.2 (3.7)	0.0 (0 – 100)	0.034
Diarrhoea²	9.5 (20.6)	0.00 (0 – 100)	6.1 (2.1)	0.0. (0 – 100)	0.140
Financial²	17.3 (27.6)	17.3 (0 – 100)	31.6 (4.3)	33.3 (0 – 100)	0.009
Global Health¹	61.9 (22.1)	61.9 (0 – 100)	52.4 (2.1)	50.0 (16.7 –100)	0.003

10.3 Brain Cancer Module: Brain tumour patients vs. spinal surgery controls

The scores on the QLQ-BN20 scales were compared between the brain tumour and spinal surgery groups using the Mann-Whitney test. The descriptive data and p-values for comparison on each scale are shown in Table 10.2. The two groups did not differ significantly on the future uncertainty scale, U = 3810.500, z = -0.230, p = 0.818. There was also no significant difference between the two groups on the visual disorder scale, U = 3459.000, z = -1.592, p = 0.111. On the single items in the questionnaire, the two groups did not differ significantly on the daytime drowsiness item, U= 3670.000, z = -0.699, p = 0.484; the itching item, U = 3737.500, z = -0.633, p = 0.527; or the bladder control item, U = 3535.000, z = -1.595, p = 0.111.

However, the spinal surgery group reported significantly higher levels of motor dysfunction than the brain tumour group, U = 3018.500, z = -2.601, p = 0.009 and also reported a higher incidence of weakness of both legs than the brain tumour group, U = 3060.500, z = -3.082, p = 0.002. The brain tumour group had higher scores on the communication deficit scale than the spinal surgery patients, U = 3011.000, z = -2.746, p = 0.006 and also had higher scores on the following single items: headache, U = 2880.000, z = -3.180, p = 0.001; seizures, U = 3321.000, z = -3.269, p = 0.001; and hair loss, U = 3564.000, z = -2.113, p = 0.035.

Table 10.2. Mean scores and comparisons of the brain tumour and spinal control groups on the QLQ-BN20 scales. p-value from Mann-Whitney test. Higher scores indicate greater symptom burden.

	Brain Tumour		**Spinal Surgery**		
	Mean (SD)	**Median (range)**	**Mean (SD)**	**Median (range)**	**p value for difference**
Future Uncertainty	35.7 (22.7)	33.3 (0 – 100)	35.3 (21.3)	33.3 (0 – 100)	0.818
Visual Disorder	10.3 (18.3)	0.0 (0 – 77.8)	6.1 (13.4)	0.0 (0 – 66.7)	0.111
Motor Dysfunction	18.8 (21.1)	11.1 (0 – 100)	27.8 (24.6)	22.3 (0 – 100)	0.009
Communication Deficit	18.0 (22.0)	11.1 (0 – 100)	9.4 (14.9)	22.3 (0 – 100)	0.006
Headache	36.0 (35.8)	33.3 (0 – 100)	19.0 (25.6)	0.0 (0 – 100)	0.001
Seizures	8.0 (22.2)	0.0 (0 – 100)	0.4 (3.8)	0.0 (0 – 33.3)	0.001
Daytime Drowsiness	24.8 (26.5)	33.3 (0 – 100)	26.4 (24.4)	33.3 (0 – 100)	0.484
Hair Loss	5.0 (15.9)	0.0 (0 – 100)	0.9 (5.3)	0.0 (0 – 33.3)	0.035

Itching	8.6 (19.8)	0.0 (0 – 100)	12.1 (26.4)	0.0 (0 – 100)	0.527
Weakness of Both Legs	8.3 (18.5)	0.0 (0 – 100)	21.2 (31.0)	0.0 (0 – 100)	0.002
Bladder Weakness	9.2 (20.6)	0.0 (0 – 100)	5.2 (17.2)	0.0 (0 – 100)	0.111

10.4 Quality of Life and Inspection Time

Since the scores for several scales on the EORTC QLQ-C30 were found to be positively or negatively skewed, non-parametric correlations (Spearman's rho) of total baseline inspection time scores and scores on each EORTC QLQ-C30 scale were carried out to determine whether specific aspects of Quality of Life were related to inspection time performance in the brain tumour and spinal surgery groups. The correlations are shown in Table 10.3.

In the brain tumour group, poorer inspection time scores were significantly correlated with poorer physical function scores, rho(n = 98) = 0.374, $p < 0.001$; poorer role functioning scores, rho(n = 98) = 0.264, p = 0.009 and poorer social functioning scores, rho(n = 98) = 0.206, p = 0.042. There were no significant correlations between any of the symptom scale scores and inspection time performance in the brain tumour group (see Table 11.3). Inspection time scores were positively correlated with global health scale scores in the brain tumour group, rho(n = 98) = 0.264, p = 0.009. Patients with higher total inspection time scores were more likely to have better self-perceived global health scores on the questionnaire.

Conversely, there were no significant correlations between total baseline inspection time scores and any of the physical, symptom or global health scales in the spinal surgery control group (see Table 10.3).

10.5 Brain Cancer Module and Inspection Time

Non-parametric correlations of baseline total inspection time scores and each scale/item on the QLQ-BN20 were also carried out for both the brain tumour and spinal surgery groups. The correlations are shown in Table 10.4.

In the brain tumour group, the motor dysfunction scale was significantly correlated with total inspection time score, rho(n = 98) = -0.275, p = 0.006. Patients who

reported greater motor impairment on the QLQ-BN20 were significantly more likely to have poorer inspection time scores. Inspection time was also significantly correlated with the 'headache' item in the brain tumour group, rho(n = 98) = 0.211, p = 0.037. Patients who reported greater effects of headache were more likely to have higher inspection time scores at baseline. No other significant correlations between baseline inspection time total score and any of the other QLQ-BN20 scales were obtained.

In the spinal surgery group, scores on the visual disorder scale were significantly correlated with baseline inspection time score, rho(n = 76) = -0.336, p = 0.003. Patients who reported higher levels of visual disorder were more likely to have lower (poorer) total inspection time scores at baseline. There was also a significant correlation between scores on the 'itching' item and baseline inspection time scores, rho(n = 76) = -0.304, p = 0.008. Patients in the spinal surgery group who had higher ratings on the itching item were more likely to have lower (poorer) inspection time scores. There were no other significant correlations between total scores on inspection time at baseline and any of the other QLQ-BN20 scales.

Table 10.3. Non-parametric (spearman's rho) correlations between total baseline inspection time scores and the functional, symptom and global health scales of the EORTC QLQ-C30. * denotes significant correlation.

	Brain Tumour Group		Spinal Surgery Control Group	
	Correlation coefficient (with baseline total inspection time score)	**p value for significance**	**Correlation coefficient (with baseline total inspection time score)**	**p value for significance**
Physical Function	0.374	< 0.001*	0.142	0.221
Role Function	0.264	0.009*	-0.158	0.173
Emotional Function	-0.005	0.958	-0.209	0.069
Cognitive Function	0.096	0.347	0.216	0.060
Social Function	0.206	0.042*	-0.189	0.102
Fatigue	-0.196	0.054	-0.073	0.529
Nausea/Vomiting	0.037	0.721	0.030	0.796
Pain	0.208	0.040*	0.055	0.635
Dyspnoea	-0.073	0.472	0.139	0.231
Insomnia	-0.060	0.556	0.015	0.895
Appetite Loss	0.050	0.623	-0.028	0.810
Constipation	-0.003	0.976	-0.008	0.947
Diarrhoea	0.047	0.647	-0.170	0.142
Financial Difficulties	-0.163	0.110	0.161	0.165
Global Health	0.264	0.009*	-0.100	0.392

Table 10.4. Non-parametric (spearman's rho) correlations between total baseline inspection time scores and the symptom scales/items on the QLQ-BN20. * denotes significant correlation.

	Brain Tumour Group		Spinal Surgery Control Group	
	Correlation coefficient (with baseline total inspection time score)	**P value for significance**	**Correlation coefficient (with baseline total inspection time score)**	**P value for significance**
Future Uncertainty	-0.016	0.877	0.007	0.950
Visual Disorder	0.022	0.830	-0.336	0.003*
Motor Dysfunction	-0.275	0.006*	-0.081	0.487
Communication Deficit	0.066	0.517	-0.086	0.459
Headache	0.211	0.037*	0.018	0.875
Seizures	0.121	0.235	0.198	0.087
Daytime Drowsiness	-0.163	0.109	-0.072	0.536
Hair Loss	0.034	0.740	-.0.096	0.411
Itching	0.036	0.723	-0.304	0.008*
Weakness of Both Legs	<0.001	0.997	<0.001	0.997
Bladder Control	-0.050	0.628	0.049	0.677

10.6 Discussion

10.6.1 Brain Tumour vs. Spinal Surgery Patients

Comparing the brain tumour patient group with the matched group of spinal surgery control patients revealed significant differences on a number of the EORTC QLQ-C30 scales. The spinal surgery group had significantly worse physical, role and social function than the brain tumour patients. They also reported significantly more fatigue, pain, insomnia, appetite loss, constipation and financial difficulties. The spinal surgery group also reported significantly poorer global health than the brain tumour patient group. There were also a number of significant differences between scores on the QLQ-BN20 in the two groups. The spinal surgery patients reported significantly more motor dysfunction and weakness of both legs on this scale. The brain tumour patients reported significantly more communication deficit, headache, seizures and concerns about hair loss than the spinal surgery group.

Several studies have assessed health-related quality of life (QoL) using a number of different self-report questionnaires in brain tumour patients. The majority of studies focus specifically on patients with high-grade gliomas and examine the changes in QoL throughout the course of the disease journey, and as a measure of the comparative effects of different treatments (Taphoorn et al., 2005). However, few studies have examined QoL in brain tumour patients by comparison with other patient groups. That the spinal surgery patient group had significantly worse physical, role and social function compared with the brain tumour patients is somewhat unexpected given the relative severity of disease in the latter group. However, this may be explained by the fact that the spinal surgery cohort may have been more physically symptomatic given the disabling effects of their disease. The majority of the brain tumour cohort likely had minimal physical impairment due to the ameliorating effects of dexamethasone. The physical effects of degenerative spinal disease likely resulted in poorer role and social functioning. Moreover, given that the spinal patients were more likely to have experienced symptoms over a

significantly longer period of time than the brain tumour cohort who tended to present on a much more urgent basis, this may explain the greater financial difficulties in the spinal surgery group given that they may have been unable to work for a longer period of time than the brain tumour patients. Furthermore, the increased pain in the spinal surgery group is unsurprising and this may also explain the increased levels of insomnia and fatigue due to pain. Again, the increased reports of motor dysfunction and weakness in both legs in the QLQ-BN20 in the spinal surgery group are common symptoms in this group of patients.

The brain tumour group had a higher symptom level on only the communication deficit, headache, seizures and hair loss items on the QLQ-BN20, suggesting that QoL was no worse in this patient group than in the matched spinal surgery cohort. In one of the few studies to compare the QoL of brain tumour patients with another patient group, Giovagnoli (1999) compared a group of glioma patients with stable disease following surgery, radiotherapy and chemotherapy with a group of control patients who had other chronic neurological disease. All patients completed the Functional Living Index-Cancer (FLIC) as a measure of QoL and there were no significant differences between the two groups on any of the FLIC scales. Therefore, although the brain tumour patients were at a different stage of disease and a different measure of QoL was used, the finding of the present study that suggest the brain tumour patients did not have significantly worse QoL than matched surgical control patients is supported by Giovagnoli (1999).

It must again be considered that the group of brain tumour patients who were entered into the study were a select cohort who likely had less severe symptoms than those who declined to participate, or were excluded for medical reasons. Therefore, although the brain tumour cohort did not appear to have a poorer health-related QoL by comparison with the spinal surgery patient group, this may not reflect the entire newly-presenting group of patients.

10.6.2 Quality of Life and Inspection Time Scores

In the brain tumour group, a number of EORTC QLQ-C30 scales were correlated with inspection time scores. Patients with poorer physical function, role function and social function were all significantly more likely to have lower (poorer) inspection time total scores. Increased pain and lower (poorer) global health score were also significantly associated with poorer inspection time total scores in the brain tumour group. Higher levels of motor dysfunction, as assessed by the QLQ-BN20, were also significantly associated with poorer inspection time scores in the brain tumour group. Interestingly, greater symptoms of headache were associated with higher (better) inspection time scores in the brain tumour group. Conversely, there were no significant correlations between any of the scales on the EORTC QLQ-C30 and total inspection time scores in the spinal surgery control group. The visual disorder and itching items in the QLQ-BN20 were, however, correlated with total baseline inspection time scores in the spinal surgery group, with greater reported symptoms associated with poorer inspection time scores.

Correlating the different scales on the two QoL questionnaires was carried out in order to determine whether, in addition to providing information regarding the presence and extent of visual information processing slowing that reflects the extent cognitive impairment, inspection time performance also reflects specific aspects of quality of life. It is well recognised that impaired cognition can have a significantly detrimental effect on an individual's quality of life (Weitzner and Meyers, 1997) and this study gives an insight into the specific areas of quality of life that may be affected by slowed information processing in brain tumour patients.

The brain tumour and spinal surgery groups differed with respect to the quality of life scales that were significantly related to total baseline inspection time scores, with poorer physical, role and social functioning reflected in poorer inspection time scores in the brain tumour group. This may be the result of increased levels of depression resulting from impairment in these aspects of daily function, since previous studies have reported slowed information processing in unmedicated depressed patients, as measured by the visual inspection time task (Tsourtos et al., 2002). However, given

that the spinal surgery group actually had significantly poorer physical, role and social function by comparison with the brain tumour group and that levels of anxiety and depression, as measured by the Hospital Anxiety and Depression Scale, did not differ, depression is unlikely to explain these correlations in the brain tumour group.

Motor dysfunction was significantly associated with poorer inspection time scores in the brain tumour group, but not in the spinal surgery group. Although a small minority of patients who experienced motor weakness as a result of the tumour location chose to voice their responses, with the researcher recording the responses, the majority of patients were able to use the computer keyboard to respond independently. However, it may be that those patients with some form of motor weakness were also more likely to experience more severe cognitive symptoms. Conversely, motor dysfunction in the spinal surgery group was related only to the physical effects of degenerative spinal disease and this may explain why there was no significant correlation between the motor function scale and inspection time performance in the spinal surgery cohort. Surprisingly, complaints of headache were significantly related to better inspection time scores in the brain tumour cohort and this finding is somewhat difficult to explain. It may be that those patients who complain of headache are those for whom headache is the primary symptom. In patients with more severe symptoms such as impaired physical and/or mental function, the effects of headache may seem relatively minor and are therefore rated as less severe in these patients. Those patients for whom headache is the only symptom would therefore likely be better functioning since they are relatively unaffected by other complaints and this could, in part, explain the correlation between increased report of headache and better inspection time performance.

It is also of interest that the visual disorder scale in the QLQ-BN20 was correlated with inspection time scores in the spinal surgery control group only, with greater visual disorder significantly related to poorer inspection time scores. This was despite the fact that there was no difference between the levels of visual disorder reported in the two groups. It would be expected that a higher symptom burden on this scale would be reflected in poorer scores on inspection time since it relies

heavily on the participant having intact visual function and it is unclear why this was not observed in the brain tumour cohort.

11 Conclusions and Future Work

This thesis has shown that newly-presenting patients with supratentorial brain tumours experience tumour-related cognitive impairment in a number of different domains at the time of presentation and prior to any treatment, by comparison with age and sex-matched surgical and healthy controls. Few studies have assessed cognition in patients with brain tumours prior to any surgical intervention and this thesis has given insight into the role of the tumour itself in causing impairment in cognition. Moreover, the inclusion of patients with meningiomas and metastases in the study cohort has shown the feasibility of carrying out such assessments in these patient groups and shows that cognitive impairment is common to patients with these types of tumours. This has been an understudied area to date. Surgical intervention, in the form of either biopsy or resection, results in general worsening of these deficits, although this may be transient, at least for those patients with a low-grade glioma. Initial post-operative deterioration was also observed on several cognitive measures in a matched surgical control cohort.

In this large cohort of brain tumour patients, there was a clear trend towards greater post-operative deterioration in those patients who had undergone a resection by comparison with those patients who had biopsy only, although the difference did not reach statistical significance. This was likely due to relatively high levels of attrition that occurred in the brain tumour cohort in particular. Chapters 8 and 9 show that there was no trend towards greater impairment in a single histological group on the baseline test measures, nor was there any specific histological group that was at greater risk of impairment following surgery. However, the high-grade glioma group showed greater impairment on inspection time, a measure of memory and a measure of verbal fluency in the pre-operative period. Post-operatively, the low-grade glioma group deteriorated significantly more than the high-grade glioma group on the Williams Delayed Recall Test and there was a trend towards greater post-operative deterioration in the low-grade glioma group by comparison with the high-grade glioma patients, although this did not reach significance. Inspection time testing on a second occasion post-operatively (session 3) suggested that the observed initial post-

operative deterioration in the low-grade cohort was transient, whereas this pattern of initial deterioration followed by a significant improvement on inspection time was not evident in the group of patients with a high-grade glioma. There were few significant differences in terms of test performance dependent upon the location of the tumour, by hemisphere or lobe.

Research has expanded in recent years and the cognitive effects of treatments and interventions in brain tumour patients are now being addressed in addition to measurements of neurological function and objective response on CT or MRI scanning. Thus, neurocognitive function is becoming an increasingly important outcome measure in clinical trials in neuro-oncology and as a result there is a need for short, repeatable tests that are sensitive to change and measure an important function (Weitzner and Meyers, 1997). This thesis primarily aimed to evaluate the utility of the inspection time task, a measure of visual information processing, as a clinical tool for use in neuro-oncological patients. The task was found to be a useful tool with which to measure and quantify the extent of slowing in visual information processing in brain tumour patients at the time of presentation, and as an indicator of response to treatment. It was as sensitive as a number of commonly-used standardised measures at baseline and was able to distinguish post-operative deterioration in the brain tumour group specifically from the general post-operative deterioration observed in both the brain tumour and surgical controls on other follow-up tests. Given that the task has a number of specific advantages for use in neuro-oncology: it is not confounded by motor or speech impairments; is suitable for repeated assessment and is a relatively short and simple measure, it could usefully be incorporated into neurocognitive test batteries in order to provide meaningful information that is not confounded by focal neurological deficits, emotional distress or practice effects. Moreover, minimal staff training is required for the administration of the inspection time task and it therefore could easily be administered in a clinic setting, with minimal demands on resources.

In addition to detailing the cognitive status of brain tumour patients in the pre- and post-operative period and evaluating the utility of inspection time testing in this

patient cohort, this thesis also raises important issues surrounding the potential for bias when carrying out cognitive testing with brain tumour patients. Low-grade glioma patients were over-represented in the study cohort and high-grade glioma patients were under-represented when compared with the total number of potentially eligible patients. As such, previous and future studies of cognition in brain tumour patients may actually under-estimate the extent of impairment, given that patients with high-grade gliomas tend to exhibit the most severe symptoms. Thus, the importance of recording and reporting the representative nature of the sample obtained in similar studies is highlighted (Scotland et al., 2009).

There are a number of directions for future research into the use of this measure of visual information processing in neuro-oncology. Given its proven utility as a measure in neuro-oncology that is tolerated by the majority of patients, the visual inspection time task could usefully be employed as a measure of brain changes that occur during and for many years following radiotherapy in patients with brain tumours. Inspection time testing during functional MRI scanning could give insight into the specific brain processes and networks that are disrupted as a result of the presence of a tumour. Future studies could also expand upon the methodology used in the present study to include other version of the inspection time that use an adaptive staircase procedure in order to determine whether those brain tumour patients who fail to meet validity criteria on the more commonly-used version of the test do so as a result of severe slowing to visual information processing ability, or because of reduced attentional function. Finally, given that low-grade glioma patients appeared to show greater deterioration than those patients with high-grade gliomas in the initial post-operative period, it is of particular interest to determine whether or not this deterioration is transient, as indicated by preliminary analysis of data from the small number of patients who completed testing on a second post-operative occasion. Therefore, further studies of cognition in glioma patients could evaluate inspection time changes post-operatively as described in this thesis, but could also include longer follow-up times in order to further evaluate the consistency of these initial changes.

References

ABEN, I., VERHEY, F., LOUSBERG, R., LODDER, J. & HONIG, A. (2002) Validity of the Beck Depression Inventory, Hospital Anxiety and Depression Scale, SCL-90, and Hamilton Depression Rating Scale as screening Instruments for Depression in Stroke Patients. *Psychosomatics,* 43**,** 386-393.

ABWENDER, D., SWAN, J. G., BOWERMAN, J. T. & CONNOLLY, S. W. (2001) Qualitative analysis of verbal fluency output: review and comparison of several scoring methods. *Assessment,* 8**,** 323-336.

AISEN, P. S., DAVIS, K. L., BERG, J. D., SCHAFER, K., CAMPBELL, K., THOMAS, R. G., WEINER, M. F., FARLOW, M. R., SANO, M., GRUNDMAN, M. & THAL, L. J. (2000) A randomized controlled trial of prednisone in Alzheimer's disease. *Neurology,* 54**,** 588-593.

ANDERSON, M. (1986) Inspection time and IQ in young children. *Personality and Individual Differences,* 7**,** 677-686.

ANDERSON, M., REID, C. & NELSON, J. (2001) Developmental changes in inspection time: what a difference a year makes. *Intelligence,* 29**,** 475-486.

ANDREWES, D. G., KAYE, A., MURPHY, M., HARRIS, B., AITKEN, S., PARR, C. & BATES, L. (2003) Emotional and social dysfunction in patients following surgical treatment for brain tumour. *Journal of Clinical Neuroscience,* 10**,** 428-433.

ARMSTRONG, C. L., GOLDSTEIN, B., SHERA, D., LEDAKIS, G. E. & TALLENT, E. M. (2003) The predictive value of longitudinal neuropsychologic assessment in the early detection of brain tumour recurrence. *Cancer,* 97**,** 649-656.

ARMSTRONG, C. L., HUNTER, J. V., LEDAKIS, G. E., COHEN, B., TALLENT, B. A., GOLDSTEIN, B. H., TOCHNER, Z., LUSTIG, R., JUDY, K. D., A, P., MOLLMAN, J. E., STANCZAK, E. M., JO, M. Y., THAN, T. L. & PHILLIPS, P. (2002) Late cognitive and radiographic changes related to radiotherapy. *Neurology,* 59**,** 40-48.

AVERY, T. L. (1971) Seven cases of frontal tumour with psychiatric presentation. *British Journal of Psychiatry,* 119**,** 19-23.

BALLDIN, J., GOTTFRIES, C.-G., KARLSSON, I., LINDSTET, G., LANGSTROM, G. & WALINDER, J. (1983) Dexamethasone suppression test and serum prolcatin in dementia disorders. *British Journal of Psychiatry,* 143**,** 277-281.

BARTOLOMEI, F., BOSMA, I., KLEIN, M., BAAYEN, J. C., REIJNEVELD, J. C., POSTMA, T. J., HEIMANS, J. J., VAN DIJK, B. W., DE MUNCK, J. C., DE JOHNGH, A., COVER, K. S. & STAM, C. J. (2006a) Disturbed functional connectivity in brain tumour patients: Evaluation by graph analysis of synchronization matrices. *Clinical Neuropsychology,* 117**,** 2039-2049.

BARTOLOMEI, F., BOSMA, I., KLEIN, M., BAAYEN, J. C., REIJNEVELD, J. C., POSTMA, T. J., HEIMANS, J. J., VAN DIJK, B. W., DE MUNCK, J. C., DE JONGH, A., COVER, K. S. & STAM, C. J. (2006b) How do brain tumours alter functional connectivity? A magnetoencephalography study. *Annals of Neurology,* 59**,** 128-138.

BASTIN, M. E., CARPENTER, T. K., ARMITAGE, A., SINHA, S., WARDLAW, J. M. & WHITTLE, I. R. (2006) Effects of dexamethasone on cerebral perfusion and water diffusion in patients with high-grade glioma. *American Journal of Neuroradiology,* 27**,** 402-408.

BATES, T., PELLETT, O., STOUGH, C. & G.L, M. (1994) The effects of smoking on simple and choice reaction time. 114**,** 365-378.

BEHIN, A., HOANG-XUAN, K., CARPENTIER, A. F. & DELATTRE, J. Y. (2003) Primary brain tumours in adults. *The Lancet,* 361**,** 323-331.

BEKKER, A. Y. & WEEKS, E. J. (2003) Cognitive function after anaesthesia in the elderly. *Best Practice & Research Clinical Anaesthesiology* 17**,** 259-272.

BENEDICT, R. H. B. & ZGALJARDIC, D. J. (1998) Practice effects during repeated administrations of memory test with and without alternate forms. *Journal of Clinical and Experimental Neuropsychology,* 20**,** 339-352.

BENTON, A. L. & HAMSHER, K. D. (1989) *Multilingual Aphasia Examination,* Iowa City, AJA Associates.

BERES, C. A. & BARON, A. (1981) Improved digit symbol substitution by older women as a result of extended practice. *Journal of Gerentology,* 36**,** 591-597.

BOAKE, C. & MEYERS, C. A. (1993) Brain tumor rehabilitation: Survey of clinical practice. *Archives of Physical Medicine and Rehabilitation,* 74**,** 1247.

BONNEY, K. R., ALMEIDA, O. P., FLICKER, L., DAVIES, S., CLARNETTEA, R., ANDERSON, M. & LAUTENSCHLAGER, N. T. (2006) Inspection time in non-demented older adults with mild cognitive impairment *Neuropsychologia,* 44**,** 1452-1456.

BORS, D. A., STOKES, T. L., FORRIN, B. & HODDER, S. L. (1999a) Inspection time and intelligence: practice, strategies and attention. *Intelligence,* 27**,** 111-129.

BORS, D. A., STOKES, T. L., FORRIN, B. & HODDER, S. L. (1999b) Inspection time and intelligence: practice, strategies, and attention. *Intelligence,* 27**,** 111-129.

BOSMA, I., DOUW, L., BARTOLOMEI, F., HEIMANS, J. J., VAN DIJK, B. W., POSTMA, T. J., STAM, C. J., REIJNEVELD, J. C. & KLEIN, M. (2008) Synchronized brain activity and neurocognitive function in patients with low-grade glioma: A magnetoencephalography study. *Neuro-Oncology,* 10**,** 734-744.

BOSMA, I., VOS, M. J., HEIMANS, J. J., TAPHOORN, M. J. B., AARONSON, N. K., POSTMA, T. J., VAN DER PLOEG, H. M., MULLER, M., VANDERTOP, W. P., SLOTMAN, B. J. & KLEIN, M. (2007) The course of neurocognitive functioning in high-grade glioma patients. *Neuro-Oncology*.

BOTWINICK, J., STORANDT, M., BERG, L. & BOLAND, S. (1988) Senile dementia of the Alzheimer type: subject attrition and testability in research. *Archives of Neurology,* 45**,** 493-496.

BRAFF, D. L., SACCUZZO, D. P. & GEYER, M. A. (1991) *Information processing dysfunctions in schizophrenia: studies of visual backward masking, sensorimotor gating and habituation*, Elsevier: New York.

BRAND, C. R. (1996) *The g factor*, Chichester:Wiley.

BRESSLER, S. L. (2002) Understanding cognition through large-scale cortical networks. *Current directions in psychological science,* 11**,** 58-61.

BRIGHT, P., JALDOW, E. & KOPELMAN, M. D. (2002) The National Adult Reading Test as a measure of premorbid intelligence: A comparison with estimates derived from demographic variables. *Journal of the International Neuropsychology Society,* 8**,** 847-854.

BROWN, P. D., BUCKNER, J. C., O'FALLON, J. R., ITURRIA, N. L., BROWN, C. A., O'NEILL, B. P., SCHEITHAUER, B. W., DINAPOLI, R. P., ARUSELL, R. M., CURRAN, W. J., ABRAMS, R. & SHAW, E. G. (2003) Effects of radiotherapy on cognitive function in patients with low-grade glioma measured by the Folstein Mini-Mental State Examination. *Journal of Clinical Oncology,* 13.

BROWN, P. D., BUCKNER, J. C., O'FALLON, J. R., ITURRIA, N. L., O'NEILL, B. P., BROWN, C. A., SCHEITHAUER, B. W., DINAPOLI, R. P., ARUSELL, R. M., CURRAN, W. J., ABRAMS, R. & SHAW, E. G. (2004) Importance of baseline mini-mental state examination as a prognostic factor for patients with low-grade glioma. *International Journal of Radiation Oncology Biology Physics,* 59**,** 117-125.

BURNS, N. R. & NETTELBECK, T. (2003) Inspection time in the structure of cognitive abilities: where does IT fit? *Intelligence,* 31**,** 237-255.

BURNS, N. R., NETTELBECK, T. & COOPER, C. J. (1999) Inspection time correlates with general speed of processing but not fluid ability. *Intelligence,* 27**,** 37-44.

BUTLER, P. D., DESANTI, L. A., MADDOX, J., HARKAVY-FRIEDMAN, J. M., AMADOR, X. F., GOETZ, R. R., JAVITT, D. C. & GORMAN, J. M. (2002) Visual backward-masking deficits in schizophrenia: relationship to visual pathway function and symptomatology. *Schizophrenia Research,* 59**,** 199-209.

BUTLER, P. D., HARKAVY-FRIEDMAN, J. M., AMADOR, X. F. & GORMAN, J. M. (1996) Backward masking in schizophrenia: relationship to medication status, neuropsychological functioning, and dopamine metabolism. *Society of Biological Psychiatry,* 40**,** 295-298.

CARROLL, J. B. (1993) *Human cognitive abilities: a survey of factor analytic studies*, Cambridge University Press: Cambridge.

CHATFIELD, M., BRAYNE, C. E. & MATTHEWS, F. E. (2005) A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies *Journal of Clinical Epidemiology,* 58**,** 13-19.

CHEAM, E. W. S., DOB, D. P., SKELLY, A. M. & LOCKWOOD, G. G. (1995) The effects of nitrous oxide on the performance of psychomotor tests. *Anaesthesia,* 50.

CLYDE, Z., CHATAWAY, S. J., SIGNORINI, D., GREGOR, A. & GRANT, R. (1998) Significant change in tests of neurological impairment in patients with brain tumours. *Journal of Neuro-Oncology,* 39**,** 81-90.

COHEN, M. J. & STANCZAK, D. E. (2000) On the reliability, validity, and cognitive structure of the Thurstone Word Fluency Test. *Archives of Clinical Neuropsychology,* 15**,** 267-279.

COLLIE, A., MARUFF, P., DARBY, D. G. & MCSTEPHEN, M. (2003) The effects of practice on the cognitive test performance of neurologically normal

individuals assessed at brief test-retest intervals. *Journal of the International Neuropsychology Society,* 9**,** 419-428.

CORREA, D. D., DEANGELIS, L. M., SHI, W., THALER, H. T., LIN, M. & ABREY, L. E. (2007) Cognitive functions in low-grade gliomas: disease and treatment effects. *Journal of Neuro-Oncology,* 81**,** 175-184.

CRAWFORD, J. R., DEARY, I. J., STARR, J. & WHALLEY, L. J. (2001) The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychological Medicine,* 31**,** 451-458.

CRAWFORD, J. R., STEWART, L. E., GARTHWAITE, P. E., PARKER, D. M. & BESSON, J. A. O. (1988) The relationship between demographic variables and NART performance in normal subjects. *Journal of Clinical Psychology,* 27**,** 181-182.

CRAWFORD, J. R., STEWART, L. E., PARKER, D. M., BESSON, J. A. O. & DE LACEY, G. (1989) Prediction of WAIS IQ with the National Adult Reading Test: cross-validation and extension. *British Journal of Clinical Psychology,* 28**,** 267-273.

CROWE, S. (2000) Does the letter number sequence task measure anything more than digit span? *Assessment,* 7**,** 113-117.

CURRAN, W. J., SCOTT, C. B., HORTON, J., NELSON, J. S., WEINSTEIN, A. S., FISCHBACH, J., CHANG, C. H., ROTMAN, M., ASBELL, S. O., KRISCH, R. E. & NELSON, D. F. (1993) Recursive Partitioning Analysis of Prognostic Factors in Three Radiation Therapy Oncology Group Malignant Glioma Trials. *Journal of the National Cancer Institute,* 85**,** 704-710.

DAVIS, K. L., BM, D., GREENWALD, B., MOHS, R., MATHE, A., JOHNS, C. & HORWATH, T. (1986) Cortisol and Alzheimer's disease: I. basal studies. *American Journal of Psychiatry,* 143**,** 300-305.

DEANGELIS, L. M. (2001) Brain Tumors. *New England Journal of Medicine,* 344**,** 114-123.

DEARY, I., J, HUNTER, R., LANGAN, S. J. & GOODWIN, G. M. (1991) Inspection time, psychometric intelligence and clinical estimates of cognitive ability in pre-senile Alzheimer's disease and Korsakoff's psychosis. *Brain,* 114**,** 2543-2554.

DEARY, I., J, SIMONONTTO, E., MARSHALL, A., MARSHALL, I., GODDARD, N. & WARDLAW, J. M. (2001) The functional anatomy of inspection time: a pilot fMRI study. *Intelligence,* 29**,** 497-510.

DEARY, I., J, SIMONONTTO, E., MEYER, M., MARSHALL, A., MARSHALL, I., GODDARD, N. & WARDLAW, J., M (2004) The functional anatomy of inspection time: an event-related fMRI study. *NeuroImage,* 22**,** 1466-1479.

DEARY, I., J & STOUGH, C. (1996) Intelligence and inspection time: achievements, prospects, and problems. *American Psychologist,* 51**,** 599-608.

DEARY, I. J. (1993) Inspection time and WAIS-R IQ subtypes: A confirmatory factor analysis study. *Intelligence,* 17**,** 223-236.

DEARY, I. J. (1995) Auditory Inspection Time and Intelligence: What Is the Direction of Causation? *Developmental Psychology,* 31**,** 237-250.

DEARY, I. J., CARYL, P. G., EGAN, V. & WIGHT, D. (1989) Visual and auditory inspection time: Their interrelationship and correlations with IQ in high ability subjects. *Personality and Individual Differences,* 10**,** 525-533.

DER, G. & DEARY, I. J. (2006) Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health and Lifestyle Survey. *Psychology and Ageing,* 21**,** 62-73.

DUFFAU, H., CAPELLE, L., DENVIL, D., SICHEZ, N., GATIGNOL, P., LOPES, M., MITCHELL, M.-C., SICHEZ, J.-P. & VAN EFFENTERRE, R. (2003) Functional recovery after surgical resection of low grade gliomas in eloquent brain: hypothesis of brain compensation. *Journal of Neurology, Neurosurgery and Psychiatry,* 74**,** 901-907.

DUNCAN, J., SEITZ, J., KOLODNY, J., BOR, D., HERZOG, H., AHMED, A., NEWELL, F. N. & EMSLIE, H. (2000) A neural basis for general intelligence. *Science,* 289**,** 457-460.

EDMONDS, C. J., ISAACS, E. B., VISSCHER, P. M., ROGERS, M., LANIGAN, J., SINGHAL, A., LUCAS, A., GRINGRAS, P., DENTON, J. & DEARY, I. J. (2008) Inspection time and cognitive abilities in twins aged 7 to 17 years: Age-related changes, heritability and genetic covariance. *Intelligence,* 36**,** 210-225.

EGAN, V. (1994) Intelligence, inspection time and cognitive strategies. *British Journal of Psychology,* 85**,** 305-315.

EK, L., SMITS, A., PAHLSON, A. & ALMKVIST, O. (2005) Analysis of cognitive dysfunction in patients with low-grade glioma. *Journal of Clinical Psychology in Medical Settings,* 12**,** 165-173.

ERBER, J. T., BOTWINICK, J. & STORANDT, M. (1981) The impact of memory on age differences in Digit Symbol performance. *Journal of Gerontology,* 36**,** 586-590.

ESPOSITO, G., KIRKBY, B. S., VAN HORN, J. D., ELLMORE, T. M. & BERMAN, K. F. (1999) Context-dependent, neural system-specific neurophysiological concomitants of ageing: mapping PET correlates during cognitive activation. *Brain,* 122**,** 963-979.

EVANS, G. & NETTELBECK, T. (1993) Inspection time: A flash mask to reduce apparent movement effects. *Personality and Individual Differences,* 15**,** 91-94.

FAYERS, P. M., AARONSON, N. K., BJORDAL, K., GROENVOLD, M., CURRAN, D. & BOTTOMLEY, A. (2001) *The EORTC QLQ-C30 Scoring Manual (3rd Edition)*, European Organisation for Research and Treatment of Cancer, Brussels 2001.

FEDER, M., RING, H. & SOLZI, P. (1989) Rehabilitation outcomes following craniotomy for intracranial meningiomas. *Journal of Neuropsychological Rehabilitation,* 3**,** 15-17.

FLICK, S. N. (1988) Managing attrition in clinical research. *Clinical Psychology Review,* 8**,** 499-515.

FLYNN, J. R. (1999) Searching for justice: The discovery of IQ gains over time. *American Psychologist,* 54**,** 5-20.

FULLER, G. N. (2008) The WHO classification of tumours of the central nervous system, 4th edition. *Archives of Pathology and Laboratory Medicine,* 132.

GALTON, F. (1883) *Inquiries into human faculty and its development,* London, Macmillan.

GEHRING, K., SITSKOORN, M., AARONSON, N. K. & TAPHOORN, M. J. B. (2008) Interventions for cognitive deficits in adults with brain tumours. *Lancet Neurology,* 7**,** 548-560.

GILROY, J. (2000) *Basic Neurology (3rd ed),* New York, McGraw-Hill Health Professions Division.

GIOVAGNOLI, A. R. (1999) Quality of life in patients with stable disease after surgery, radiotherapy and chemotherapy for malignant brain tumour. *Journal of Neurology, Neurosurgery and Psychiatry,* 67**,** 358-363.

GOLDSTEIN, B., ARMSTRONG, C., L., JOHN, C. & TALLENT, E., M. (2003) Attention in adult intracranial tumors patients. *Journal of Clinical and Experimental Neuropsychology,* 25**,** 66-78.

GRANT, R. (2004) Overview: brain tumour diagnosis and management/ Royal College of Physicians guidelines. *Journal of Neurology, Neurosurgery, and Psychiatry,* 75**,** 18-23.

GRANT, R., SLATTERY, J., GREGOR, A. & WHITTLE, I. R. (1994) Recording neurological impairment in clinical trials of glioma. *Journal of Neuro-Oncology,* 19**,** 37-49.

GREGOR, A., CULL, A., TRAYNOR, E., STEWART, M., LANDER, F. & LOVE, S. (1996) Neuropsychometric evaluation of long-term survivors of adult brain tumours: relationship with tumour and treatment parameters *Radiotherapy and Oncology,* 41**,** 55-59.

GREGORY, T., NETTELBECK, T. & WILSON, C. (2009) Within-person changes in inspection time predict memory. *Personality and Individual Differences,* 46**,** 741-743.

GRUDNIK, J. L. & KRANZLER, J. H. (2001) Meta-analysis of the relationship between intelligence and inspection time *Intelligence,* 29**,** 523-535.

HAALAND, K. Y. & DELANEY, H. D. (1981) Motor deficits after left or right hemisphere damage due to stroke or tumor. *Neuropsychologia,* 19**,** 17-27.

HABETS, E. J. J., WALCHENBACK, R., KLOET, A., VECHT, C. J., ZWINKELS, H. & TAPHOORN, M. J. B. (2008) Improvement of cognitive functioning following surgery in brain tumour patients. *8th meeting of the European Association for Neuro-Oncology.* Barcelona, Spain.

HAHN, C. A., DUNN, R. H., LOGUE, P. E., KING, J. H., EDWARDS, C. L. & HALPERIN, E. C. (2003) Prospective study of neuropsychologic testing and quality-of-life assessment of adults with primary malignant brain tumours. *International Journal of Radiation Oncology Biology Physics,* 55**,** 992-999.

HALL, W. A. (1998) The safety and efficacy of stereotactic biopsy for intracranial lesions. *American Cancer Society,* 82**,** 1749-1755.

HART, M. G., GRANT, R. & METCALFE, S. E. (2000) Biopsy versus resection for high grade glioma. *Cochrane Database of Systematic Reviews*.

HERMAN, M. A., TREMONT-LUKATS, I., MEYERS, C. A., TRASK, D. D., FROSETH, C., RENSCHLER, M. F. & MEHTA, M. P. (2003) Neurocognitive and functional assessment of patients with brain metastases. A pilot study. *American Journal of Clinical Oncology,* 26**,** 273-279.

HIRSHORN, E. A. & THOMPSON-SCHILL, A. L. (2006) Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency. *Neuropsychologia,* 44**,** 2547-2557.

HOGERVORST, E., COMBRINCK, M., LAPUERTA, P., RUE, J., SWALES, K. & BUDGE, M. (2002) The Hopkins Verbal Learning Test and screening for dementia. *Dementia and geriatric Cognitive Disorders,* 13**,** 13-20.

HOM, J. & REITAN, R. M. (1984) Neuropsychological correlates of rapidly vs slowly growing intrinsic cerebral neoplasms. *Journal of Clinical Neuropsychology,* 6**,** 309-324.

HOWE, M. J. A. (1988) Intelligence as an explanation. *British Journal of Psychology,* 79**,** 349-360.

IRWIN, R. J. (1984) Inspection time and its relation to intelligence. *Intelligence,* 8**,** 47-65.

ISAT, I. S. A. T. I. C. G. (2002) International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomized trial. *Lancet* 360**,** 1267-1274.

JEREMIC, B., GRUJICIC, D., ANTUNOVIC, V., DJURIC, L., STOJANOVIC, M. & SHIBAMOTO, Y. (1994) Influence of extent of surgery and tumor location on treatment outcome of patients with glioblastoma multiforme treated with combined modality approach. *Journal of Neuro-Oncology,* 21**,** 177-185.

JOHNSON, A. M., ALMEIDA, Q. J., STOUGH, C., THOMPSON, J. C., SINGARAYER, R. & JOG, M. S. (2004a) Visual inspection time in Parkinson's disease: deficits in early stages on cognitive processing. *Neuropsychologia,* 42**,** 577-583.

JOHNSON, W., BOUCHARD JR, T. J., KRUEGER, R. F., MCGUE, M. & GOTTESMAN, I. I. (2004b) Just one g: Consistent results from three test batteries. *Intelligence,* 32**,** 95-107.

JOHNSON, W., TE NIJENHUIS & T.J, B. J. (2008) Still just 1 g: Consistent results from five test batteries *Intelligence,* 36**,** 81-95.

JOKEIT, H. & EBNER, A. (1999) Long term effects of refractory temporal lobe epilepsy on cognitive abilities: a cross sectional study. *Journal of Neurology, Neurosurgery and Psychiatry,* 67**,** 44-50.

KALMIJN, S., LAUNER, L. J., STOLK, R. P., DE JONG, F. H., POLS, H. A., HOFMAN, A., BRETELER, M. M. & LAMBERTS, S. W. (1998) A prospective study on cortisol, dehydroepiandrosterone sulfate and cognitive function in the elderly. *Journal of Clinical Endocrinology and Metabolism,* 83**,** 3487-3492.

KARNOFSKY, D. A. & BURCHENAL, J. H. (1943) The clinical evaluation of chemotherapeutic agents in cancer. *Evaluation of Chemotherapeutic Agents.* New York, Columbia University Press.

KAYL, A. E. & MEYERS, C. A. (2003) Does brain tumour histology influence cognitive function? *Neuro-Oncology,* 5**,** 255-260.

KERSHNER, P. & WANG-CHENG, R. (1989) Psychiatric side effects of steroid therapy. *Psychosomatics,* 30**,** 135-139.

KILBRIDE, L., SMITH, G. & GRANT, R. (2007) The frequency and cause of anxiety and depression amongst patients with malignant brain tumours between surgery and radiotherapy. *Journal of Neuro-Oncology,* 84**,** 297-304.

KIRSCHBAUM, C., WOLF, O. T., MAY, M., WIPPICH, W. & HELLHAMMER, D. H. (1996) Stress- and treatment-induced elevations of cortisol levels

associated with impaired declarative memory in healthy adults. *Life Sciences,* 58**,** 1475-1483.

KLEIN, M., ENGELBERTS, N. H. J., VAN DER PLOEG, H. M., KASTELEIJN-NOLST TRENITE, D. G. A., AARONSON, N. K., TAPHOORN, M. J. B., BAAIJEN, H., VANDERTOP, W. P., MULLER, M., POSTMA, T. J. & HEIMANS, J. J. (2003a) Epilepsy in low-grade gliomas: the impact on cognitive function and quality of life. *Annals of Neurology,* 54**,** 514-520.

KLEIN, M. & HEIMANS, J. J. (2004) The measurement of cognitive functioning in low-grade glioma patients after radiotherapy. *Journal of Clinical Oncology***,** 966-967.

KLEIN, M., HEIMANS, J. J., AARONSON, N. K., VAN DER PLOEG, H. M., GRIT, J., MULLER, M., POSTMA, T. J., MOOIJ, J. J., BOERMAN, R. H., BEUTE, G. N., OSSENKOPPELE, G. J., VAN IMHOFF, G. W., DEKKER, A. W., JOLLES, J., SLOTMAN, B. J., STRUIKMANS, H. & TAPHOORN, M. J. B. (2002) Effect of radiotherapy and other treatment-related factors on mid-term to long-term cognitive sequelae in low-grade gliomas: a comparative study. *The Lancet* 360**,** 1361-1368.

KLEIN, M., POSTMA, T. J., TAPHOORN, M. J. B., AARONSON, N. K., VANDERTOP, W. P., MULLER, M., VAN DER PLOEG, H. M. & HEIMANS, J. J. (2003b) The prognostic value of cognitive functioning in the survival of patients with high-grade glioma. *Neurology,* 61**,** 1796-1798.

KLEIN, M., TAPHOORN, M. J. B., HEIMANS, J. J., VAN DER PLOEG, H. M., VANDERTOP, W. P., SMIT, E. F., LEENSTRA, S., TULLEKEN, C. A. F., BOOGERD, W., BELDERBOS, J. S. A., CLEIJNE, W. & AARONSON, N. K. (2001) Neurobehavioural status and health-related quality of life in newly diagnosed high-grade glioma patients. *Journal of Clinical Oncology,* 19**,** 4037-4047.

KOEHLER, P. J. (1995) Use of corticosteroids in neuro-oncology. *Anticancer Drugs,* 6**,** 19-33.

KORTE, K. B., HORNER, M. D. & WIDHAM, W. K. (2002) The Trail Making Test, Part B: cognitive flexibility of ability to maintain set? *Applied Neuropsychology,* 9**,** 106-109.

KRANZLER, J. H. & JENSEN, A. R. (1989) Inspection time and intelligence: a meta-analysis. *Intelligence,* 13**,** 329-347.

KRETH, F. W., WARNKE, P. C., SCHEREMET, R. & OSTERTAG, C. B. (1993) Surgical resection and radiation therapy in the treatment of glioblastoma multiforme. *Journal of Neurosurgery,* 78**,** 762-766.

KWAN, P. & BRODIE, M. J. (2001) Neuropsychological effects of epilepsy and antiepileptic drugs. *The Lancet,* 357**,** 216-222.

LACY, M. A., GORE JR., P. A., PLISKIN, N. H., HENRY, G. K., HEILBRONNER, R. L. & HAMER, D. P. (1996) Verbal fluency task equivalence. *The Clinical Neuropsychologist,* 10**,** 305-308.

LAURSEN, P. (1997) The impact of aging on cognitive functions. An 11 year follow-up study of four age cohorts. *Acta Neurologica Scandinavica. Supplementum,* 172**,** 7-86.

LEININGER, B. E., GRAMLING, S. E., FARRELL, A. D., KREUTZER, J. S. & PECK, E. A. (1990) Neuropsychological deficits in symptomatic minor head

injury patients after concussion and mild concussion. *Journal of Neurology, Neurosurgery and Psychiatry,* 53**,** 293-296.

LEVIN, B. E., KATZEN, H. L., KLEIN, B. & LLABRE, M. L. (2000) Cognitive Decline Affects Subject Attrition in Longitudinal Research *Journal of Clinical and Experimental Neuropsychology,* 22**,** 580-586.

LEZAK, M. D., HOWIESON, D. B. & LORING, D. W. (2004) *Neuropsychological Assessment,* New York, Oxford University Press.

LUBIN, M. & FERNANDEZ, J. M. (1986) The relationship between psychometric intelligence and inspection time. *Personality and Individual Differences,* 7**,** 653-657.

LUCIANO, M., POSTHUMA, D., WRIGHT, M. J., DE GEUS, E. J. C., SMITH, G. A., GEFFEN, G. M., BOOMSMA, D. I. & MARTIN, N. G. (2005) Perceptual speed does not cause intelligence, and intelligence does not cause perceptual speed. *Biological Psychology,* 70**,** 1-8.

LUCIANO, M., SMITH, G. A., WRIGHT, M. J., GEFFEN, G. M., GEFFEN, L. B. & MARTIN, N. G. (2001) On the heritability of inspection time and its covariance with IQ: a twin study. *Intelligence,* 29**,** 443-457.

LUCIANO, M., WRIGHT, M. J., GEFFEN, G. M., GEFFEN, L. B., SMITH, G. A. & MARTIN, N. G. (2004) A genetic investigation of the covariance among inspection time, choice reaction time, and IQ subtest scores. *Behaviour Genetics,* 34**,** 41-50.

LUPIEN, S., GILLIN, C. J. & HAUGER, R. L. (1999) Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose-response study in humans. *Behavioural Neuroscience,* 113**,** 420-430.

MACKINTOSH, N. J. (1998) *IQ and human intelligence*, OUP: Oxford.

MAHONEY, F. I. & BARTHEL, D. W. (1965) Functional evaluation: The Barthel Index. *Maryland State Medical Journal,* 14**,** 61-65.

MATARAZZO, J. D. & HERMAN, D. O. (1984) Base rate data for the WAIS-R: Test-retest stability and VIQ-PIQ differences. *Journal of Clinical Neuropsychology,* 6**,** 351-366.

MATTHEWS, F. E., CHATFIELD, M., FREEMAN, C., MCCRACKEN, C., BRAYNE, C. & CFAS, M. (2004) Attrition and bias in the MRC cognitive function and ageing study: an epidemiological investigation. *BMC Public Health,* 4**,** 12.

MCGURN, M. B., STARR, J. M., TOPFER, J. A., PATTIE, A., WHITEMAN, M. C., LEMMON, H. A., WHALLEY, L. J. & DEARY, I. J. (2004) Pronunciation of irregular words is preserved in dementia, validating premorbid estimation. *Neurology,* 62**,** 1184-1186.

MCKINNEY, P. A. (2004) Brain tumours: Incidence, survival, and aetiology. *Journal of Neurology, Neurosurgery and Psychiatry,* 75**,** ii12 - ii17.

MEYERS, C., A & HESS, K., R (2003) Multifaceted end points in brain tumor clinical trials: Cognitive deterioration precedes MRI progression. *Neuro-Oncology,* 5**,** 89-95.

MEYERS, C. A. & BROWN, P. D. (2006) Role and relevance of neurocognitive assessment in clinical trials of patients with CNS tumors. *Journal of Clinical Oncology,* 24**,** 1305-1309.

MEYERS, C. A., HESS, K. R., YUNG, W. K. A. & LEVIN, V. A. (2000) Cognitive function as a predictor of survival in patients with recurrent malignant glioma. *Journal of Clinical Oncology,* 18**,** 646-650.

MEYERS, C. A. & WEFEL, J. S. (2003) The use of the Mini-Mental State Examination to assess cognitive functioning in cancer trials: No ifs, ands, buts, or sensitivity. *Journal of Clinical Oncology,* 19**,** 3557-3558.

MICELI, G., CALTAGIRONE, C., GAINOTTI, G., MASULLO, C. & SILVERI, M. C. (1981) Neuropsychological correlates of localized cerebral lesions in nonaphasic brain-damaged patients. *Journal of Clinical Neuropsychology,* 3**,** 53-63.

MIHELIC, A. H. & CRIMMINS, E. M. (1997) Loss to follow-up in a sample of Americans 70 years of age and older: The LSOA 1984-1990. *Journal of Gerontology,* 52B**,** s37 - s48.

MINDER, C. E., MULLER, T., GILLMAN, G., BECK, J. C. & STUCK, A. E. (2002) Subgroups of refusers in a disability prevention trial in older adults: baseline and follow-up analysis. *Americal Journal of Public Health,* 92**,** 445-450.

MOLLER, I. T. (1997) Cerebral dysfunction after anaesthesia. *Acta Anaesthesiologica Scandinavica,* 110**,** 13-16.

NAKAHACHI, T., ISHII, R., IWASE, M., CANUET, L., TAKAHASHI, H., KURIMOTO, R., IKEZAWA, K., AZECHI, M., SEKIYAMA, R., HONAGA, E., UCHIUMI, C., IWAKIRI, M., MOTOMURA, N. & TAKEDA, M. (2008) Frontal activity during the digit symbol substitution test determined by multichannel near-infrared spectroscopy. *Neuropsychobiology,* 57**,** 151-158.

NETTELBECK, T. (2001) Correlation between inspection time and psychometric abilities. A personal interpretation. *Intelligence,* 29**,** 459-474.

NETTELBECK, T. & KIRBY, N. H. (1983) Retarded–nonretarded differences in speed of processing. *Australian Journal of Psychology* 35**,** 445-453.

NETTELBECK, T. & LALLY, M. (1976) Inspection time and measured intelligence. *British Journal of Psychology,* 67**,** 17-22.

NETTELBECK, T. & LALLY, M. (1979) Age, intelligence, and inspection time. *American Journal of Mental Deficiency,* 83**,** 398-401.

NETTELBECK, T. & RABBITT, P. M. A. (1992) Aging, cognitive performance, and mental speed. *Intelligence,* 16**,** 189-205.

NETTELBECK, T. & VITA, P. (1992) Inspection time in two childhood age cohorts: a constant or a developmental function? *British Journal of Developmental Psychology,* 10**,** 189-197.

NETTELBECK, T. & WILSON, C. (1985) A cross-sectional analysis of developmental differences in speed of visual information processing. *Journal of Experimental Child Psychology,* 40**,** 1-22.

NETTELBECK, T. & WILSON, C. (2004) The Flynn effect: Smarter not faster *Intelligence,* 32**,** 85-93.

NEWCOMER, J. W., SELLA, H. & A.K, M. (1994) Glucocorticoid-induced impairment in declarative memory performance in adult humans. *Journal of Neuroscience,* 56**,** 527-533.

NORTH, C. A., NORTH, R. B., EPSTEIN, J. A., PIANTADOSI, S. & WHARAM, M. D. (1990) Low-grade cerebral astrocytomas: Survival and quality of life after radiation therapy. *Cancer,* 66**,** 6-14.

NUSSBAUM, E. S., DJALILIAN, H. R., CHO, K. H. & HALL, W. A. (1996) Brain Metastases. Histology, multiplicity, surgery, survival. *American Cancer Society,* 78**,** 1781-1788.

O'CARROLL, R. E., PRENTICE, N., MURRAY, C., VAN BECK, M., EBMEIER, K. P. & GOODWIN, G. M. (1995) Further evidence that reading ability is not preserved in Alzheimer's disease. *British Journal of Psychiatry,* 167**,** 659-662.

OSMON, D. C. & JACKSON, R. (2002) Inspection time and IQ: Fluid or perceptual aspects of intelligence? *Intelligence,* 30**,** 119-127.

OWEN, A. M., SAHAKIAN, B. J., SEMPLE, J., POLKEY, C. E. & ROBBINS, T. W. (1995) Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia,* 33**,** 1-24.

PAHLSON, A., EK, L., AHLSTROM, G. & SMITS, A. (2003) Pitfalls in the assessment of disability in individuals with low-grade gliomas. *Journal of Neuro-Oncology,* 65**,** 149-158.

PARKER, D. M., CRAWFORD, J. R. & STEPHEN, E. (1999) Auditory inspection time and intelligence: a new spatial localisation task. *Intelligence,* 27**,** 131-139.

PARKS, R. W., LOEWENSTEIN, D. A., DODRILL, K. L., BARKER, W. W., YOSHII, F. & CHANG, J. Y. (1988) Cerebral metabolic effects of a verbal fluency test: A PET scan study. *Journal of Clinical and Experimental Neuropsychology,* 10**,** 565-575.

PATCHELL, R. A. (2003) The management of brain metastases. *Cancer Treatment Reviews,* 29**,** 533-540.

PERAUD, A., MESCHEDE, M., EISNER, W., ILMBERGER, J. & REULEN, H.-J. (2002) Surgical resection of Grade II astrocytomas in the superior frontal gyrus. *Neurosurgery,* 50**,** 966-977.

PETRILL, S. A., LUO, D., THOMPSON, L. & DETTERMAN, D. K. (2001) Inspection time and the relationship among elementary cognitive tasks, general intelligence, and specific cognitive abilities *Intelligence,* 6**,** 487-469.

PIGNATTI, F., VAN DEN BENT, M., CURRAN, D., DEBRUYNE, C., SYLVESTER, R., THERASSE, P., AFRA, D., CORNU, P., BOLLA, M., VECHT, C. & KARIM, A., .B.M.F (2002) Prognostic factors for survival in adult patients with cerebral low-grade glioma. *Journal of Clinical Oncology,* 20**,** 2076-2084.

POSTHUMA, D., DE GEUS, E. J. C. & BOOMSMA, D. I. (2001) Perceptual speed and IQ are associated through common genetic factors. *Behaviour Genetics,* 31**,** 593-602.

PRINGLE, A.-M., TAYLOR, R. & WHITTLE, I. R. (1999) Anxiety and depression in patients with an intracranial neoplasm before and after tumour surgery. *British Journal of Neurosurgery,* 13**,** 46-51

PUGSLEY, W., KLINGER, L., PASCHALIS, C., TREASURE, T., HARRISON, M. & NEWMAN, S. (1994) The impact of microemboli during cardiopulmonary bypass on neuropsychological functioning *Stroke,* 25**,** 1393-1399.

PURCELL, R., MARUFF, P., KYRIOS, M. & PANTELIS, C. (1997) Neuropsychological function in young patients with unipolar major depression. *Psychological Medicine,* 27**,** 1277-1285.

RASMUSSEN, L. S. (1999) Postoperative cognitive decline: the extent of the problem. *Acta Anaesthesia Belgica,* 50**,** 199-204.

RASMUSSEN, L. S., LARSEN, K., HOUX, P., SKOVGAARD, L. T., HANNING, C. D., MOLLER, J. T. & GROUP, A. T. I. (2001) The assessment of postoperative cognitive dysfunction. *Acta Anaesthesiologica Scandinavica,* 45**,** 275-289.

REGINE, W. F., SCHMITT, F. A., SCOTT, C. B., DEARTH, C., PATCHELL, R. A., NICHOLS, R. C., GORE, E. M., FRANKLIN, R. L., SUH, J. H. & MEHTA, M. P. (2004) Feasibility of neurocognitive outcome evaluations in patients with brain metastases in a multi-institutional cooperative group setting: results of radiation therapy oncology group trial BR-0018. *International Journal of Radiation Oncology Biology Physics,* 58**,** 1346-1352.

REIJNEVELD, J. C., SITSKOORN, M. M., KLEIN, M., NUYEN, J. & TAPHOORN, M. J. B. (2001) Cognitive status and quality of life in patients with suspected versus proven low-grade gliomas. *Neurology,* 56**,** 618-623.

REITAN, R. M. (1958) Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills,* 8**,** 271-276.

RIJSDIJK, F. V. & BOOMSMA, D. I. (1997) Genetic mediation of the correlation between peripheral nerve conduction velocity and IQ. *Behaviour Genetics,* 27**,** 87-98.

SADLER, A. J. & DEARY, I. J. (1996) Cerebral asymmetries in inspection time. *Neuropsychologia,* 34**,** 283-295.

SALANDER, P., KARLSSON, T., BERGENHEIM, T. & HENRIKSSON, R. (1995) Long-term memory deficits in patients with malignant gliomas. *Journal of Neuro-Oncology,* 25**,** 227-238.

SALTHOUSE, T. A. (1996) The processing-speed theory of adult age differences in cognition. *Psychological Review,* 103**,** 403-428.

SAWAYA, R., HAMMOUD, M., SCHOPPA, D., HESS, K., WU, S., SHI, W.-M. & WILDRICK, D. (1998) Neurosurgical outcomes in a modern series of 400 craniotomies for treatment of parenchymal tumors. *Neurosurgery,* 42**,** 1044-1055.

SCHEIBEL, R. S., MEYERS, C., A & LEVIN, V. A. (1996) Cognitive dysfunction following surgery for intracerebral glioma: influence of histopathology, lesion location, and treatment. *Journal of Neuro-Oncology,* 30**,** 61-69.

SCHWINDACK, C., SIMINOTTO, E., MEYER, M., MCNAMARA, A., MARSHALL, I., WARDLAW, J. M. & WHITTLE, I. R. (2005) Real-time functional magnetic resonance imagine (rt-fMRI) in patients iwth brain tumours: preliminary findings using motor and language paradigms. *British Journal of Neurosurgery,* 19**,** 25-32.

SCOTLAND, J. L., AL-SHAHI SALMAN, R., DEARY, I. J. & WHITTLE, I. R. (2009) Recruitment difficulties in brain tumour patients cause participation bias: findings from a neuropsychological

study of adult inpatients with supratentorial intracranial tumours. *Acta Neurochirurgica,* online first.

SEGALOWITZ, S. J., UNSAL, A. & DYNWAN, J. (1992) CNV evidence for the distinctiveness of frontal and posterior neural processes in a traumatic brain-injured population. *Journal of Clinical and Experimental Neuropsychology,* 14**,** 545-565.

SHEA, S., BIGGER, J. T. J., CAMPION, J., FLEISS, J. L., ROLNITZKY, L. M., SCHRON, E., GORKIN, L., HANDSHAW, K., KINNEY, M. R. & BRANYON, M. (1992) Enrollment in clinical trials: institutional factors affecting enrollment in the Cardiac Arrhythmia Suppression Trail (CAST). *Controlled Clinical Trials,* 13**,** 466-486.

SHIPLEY, B. A., DEARY, I., J, TAN, J., CHRISTIE, G. & STARR, J. M. (2002) Efficiency of temporal order discrimination as an indicator of bradyphrenia in Parkinson's disease: the inspection time loop task. *Neuropsychologia,* 40**,** 1488-1493.

SIMOCA, I., OLARESCU, A. A., JIPESCU, I. & LISIEVICI, M. (1994) Postooperative outcome of intracranial meningiomas; long term prognosis. *Romanian Journal of Neurology and Psychiatry,* 32**,** 237-251.

SIMPSON, J. R., HORTON, J., SCOTT, C., CURRAN, W. J., RUBIN, P., FISCHBACH, J., ISAACSON, S., ROTMAN, M., ASBELL, S. O. & NELSON, J. S. (1993) Influence of location and extent of surgical resection on survival of patients with glioblastoma multiforme: results of three consecutive Radiation Therapy Oncology Group (RTOG) clinical trials. *International Journal of Radiation Oncology, Biology, Physics,* 26**,** 239-244.

SMITH, J. S., CHANG, E. F., LAMBORN, K. R., CHANG, S. M., PRADOS, M. D., CHA, S., TIHAN, T., VANDENBERG, S., MCDERMOTT, M. W. & BERGER, M. S. (2008) Role of extent of resection in the long-term outcome of low-grade hemispheric tumours. *Journal of Clinical Oncology,* 26**,** 1338-1345.

SPEARMAN, C. (1904) General intelligence, objectively determined and measured. *American Journal of Psychology,* 15**,** 201-293.

STAM, C. J., VAN CEPPELLEN VAN WLASUN, A. M., PIJNEBURG, Y. A., BERENDSE, H. W., DE MUNCK, J. C., SCHELTENS, P. & VAN DIJK, B. W. (2002) Generalized synchronization of MEG recordings in Alzheimer's disease: evidence for involvement of the gamma band. *Journal of Clinical Neurophysiology,* 19**,** 562-574.

STERN, P. C. & CARSTENSEN, L. L. (2000) *The aging mind: Opportunities in cognitive research*, National Research Council, National Academy Press, Washington, D.C. .

STOUGH, C., BREBNER, J., NETTELBECK, T., COOPER, C. J., BATES, T. & MANGAN, G. L. (1996) The relationship between intelligence, personality and inspection time. *British Journal of Psychology,* 87**,** 255-268.

STOUGH, C., MANGAN, G., BATES, T., KERKIN, B., PELLETT, O. & FRANK, N. (1995) Effects of nicotine on perceptual speed *Psychopharmacology,* 119**,** 305-310.

STOUGH, C., THOMPSON, J. C., BATES, T. C. & NATHAN, P. J. (2001) Examining neurochemical determinants of inspection time. Development of a biological model. *Intelligence,* 29**,** 511-522.

STRACHAN, M. W. J., DEARY, I. J., EWING, F. M. E., FERGUSON, S. S. C., YOUNG, M. J. & FRIER, B. M. (2001) Acute hypoglycaemia impairs functions of the central but not the peripheral nervous system. *Physiology & Behaviour,* 72**,** 83-92.

STUMMER, W., REULEN, H.-J., MEINEL, T., PICJLMEIER, U., SCHUMACHER, W., TONN, J.-C., ROHDE, V., OPPEL, F., TUROWSKI, B., WOICIECHOWSKY, C., FRANZ, K., PIETSCH, T. & GROUP, A.-G. S. (2008) Extent of resection and survival in glioblastoma multiforme: identification of and adjustment for bias. *Neurosurgery,* 62**,** 564-576.

STUPP, R., MASON, W. P., VAN DEN BENT, M., WELLER, M., FISHER, B., TAPHOORN, M. J. B., BELANGER, K., BRANDES, A. A., MAROSI, C., BOGDAHN, U., CURSCHMANN, J., JANZER, R. C., LUDWIN, S. K., GORLIA, T., ALLGEIER, A., LACOMBE, D., CAIRNCROSS, J. G., EISENHAUER, E. & MIRIMANOFF, R. O. (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *The New England Journal of Medicine,* 352**,** 987-996.

SWEET, J. J., NIES, K. J., LORBER, R. & VICK, N. A. (1994) Relative absence of neuropsychological deficit in patients with low-grade astrocytomas. *Journal of Clinical Psychology in Medical Settings,* 1**,** 83-104.

TAPHOORN, M. J. B. & KLEIN, M. (2004) Cognitive deficits in adult patients with brain tumours. *The Lancet Neurology,* 3**,** 159-168.

TAPHOORN, M. J. B., STUPP, R., COENS, C., OSOBA, D., KORTMANN, R., VAN DEN BENT, M. J., MASON, W., MIRIMANOFF, R. O., BAUMERT, B. G., EISENHAUER, E., FORSYTH, P. & BOTTOMLEY, A. (2005) Health-related quality of life in patients with glioblastoma: a randomised control trial. *Lancet Oncology,* 6.

TAYLOR, B. V., BUCKNER, J. C., CASCINO, T. L., O'FALLON, J. R., SCHAEFER, P. L., DINAPOLL, R. P. & SCHOMBERG, P. (1998) Effects of radiation and chemotherapy on cognitive function in patients with high-grade glioma. *Journal of Clinical Oncology,* 16**,** 2195-2201.

TEIXIDOR, P., GATIGNOL, P., LEROY, M., MASUET-AUMATELL, C., CAPELLE, L. & DUFFAU, H. (2006) Assessment of verbal working memory before and after surgery for low-grade glioma. *Journal of Neuro-Oncology,* 81**,** 305-313.

THOMSON, J. C., STOUGH, C., NATHAN, P. J., AMES, P. J. & RITCHIE, C. (2000) Effects of nicotinic antagonist mecamylamine on inspection time. *Psychopharmacology,* 150**,** 117-119.

TORRES, I. J., MUNDT, A. J., SWEENEY, P. J., LLANES-MACY, S., DUNAWAY, L., CASTILLO, M. & MACDONALD, R. L. (2003) A longitudinal neuropsychological study of partial brain radiation in adults with brain tumours. *Neurology,* 60**,** 1113-1118.

TROYER, A. K. (2000) Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology,* 22**,** 370-378.

TSOURTOS, G., THOMPSON, J. C. & STOUGH, C. (2002) Evidence of an early information processing speed deficit in unipolar major depression. *Psychological Medicine,* 32**,** 259-265.

TUCHA, O., SMELY, C., PREIER, M., BECKER, G., PAUL, G. M. & LANGE, K. W. (2003) Preoperative and postoperative cognitive functioning in patients with frontal meningiomas. *Journal of Neurosurgery,* 98**,** 21-31.

TUCHA, O., SMELY, C., PREIER, M. & LANGE, K. W. (2000) Cognitive deficits before treatment among patients with brain tumors. *Neurosurgery,* 47**,** 324-334.

VAN BEIJSTERVELDT, C. E. M., VAN BOXTEL, M. P. J., BOSMA, H., HOUX, P. J., BUNTINX, F. & JOLLES, J. (2002) Predictors of attrition in a longitudinal cognitive aging study: The Maastricht Aging Study (MAAS). *Journal of Clinical Epidemiology,* 55**,** 216-223.

VAN DEN BENT, M. (2003) The role of chemotherapy in brain metastases. *European Journal of Cancer,* 39**,** 2114-2120.

VAN DEN BENT, M. J., AFRA, D., DE WITTE, O., BEN HASSEL, M., SCHRAUB, HOANG-XUAN, K., MALMSTROM, P.-O., COLLETTE, L., PIERART, M., MIRIMANOFF, R. O., KARIM, A. B. M. F. & COUNCIL, F. T. E. R. A. B. T. G. A. T. U. M. R. (2005) Long-term efficacy of early versus delayed radiotherapy for low-grade astrocytoma and oligodendroglioma in adults: the EORTC 22845 randomised trial. *The Lancet,* 366**,** 985-990.

VECHT, C. J., AVEZAAT, C. J. J., VAN PUTTEN, W. L. J., EIJKENBOOM, W. M. H. & STEFANKO, S. Z. (1990) The influence of the extent of surgery on the neurological function and survival in malignant glioma. A retrospective analysis in 243 patients. *Journal of Neurology, Neurosurgery, and Psychiatry,* 53**,** 466-471.

VICKERS, D. (1970) Evidence for an accumulator model of psychophysical discrimination. *Ergonomics,* 13**,** 37-58.

VICKERS, D., NETTELBECK, T. & WILLSON, R. J. (1972) Perceptual indices of performance: the measurement of 'inspection time' and 'noise' in the visual system. *Perception,* 1**,** 263-295.

VICKERS, D. & SMITH, P. L. (1986) The rationale for the inspection time index. *Personality and Individual Differences,* 7**,** 609-623.

VIGLIANI, M. C., SICHEZ, N., POISSON, M. & DELATTRE, J. Y. (1996) A prospective study of cognitive functions following conventional radiotherapy for supratentorial gliomas in young adults: 4-year results. *International Journal of Radiation Oncology Biology Physics,* 35**,** 527-533.

WAITER, G. D., FOX, H. C., MURRAY, A. D., STARR, J. M., STAFF, R., T, BOURNE, V. J., WHALLEY, L. J. & DEARY, I. J. (2008) Is retaining the youthful functional anatomy underlying speed of information processing a signature of successful cognitive ageing? An event-related fMRI study of inspection time performance *NeuroImage,* 41**,** 581-595.

WECHSLER, D. (1981) *WAIS-R Manual,* New York, The Psychological Corporation.

WECHSLER, D. (1997) *Wechsler Adult Intelligence Scale-III (WAIS-III)*, The Psychological Corporation.

WEINER, M. F., VOBACH, S., OLSSON, K., SVETLIK, D. & RISSER, R. C. (1997) Cortisol secretion and Alzheimer's disease progression. *Biological Psychiatry,* 42**,** 1030-1038.

WEITZNER, M. A. & MEYERS, C., A (1997) Cognitive functioning and quality of life in malignant glioma patients: a review of the literature. *Psycho-oncology,* 6**,** 169-177.

WESSELS, P. H., WEBER, W. E. J., RAVEN, G., RAMAEKERS, F. C. S., HOPMAN, A. H. N. & TWIJNSTRA, A. (2003) Supratentorial grade II astrocytoma: biological features and clinical course. *The Lancet Neurology,* 2**,** 395-403.

WESTPHAL, M., RAM, Z., RIDDLE, V., HILT, D. & BORTEY, E. (2006) Gliadel wafer in initial surgery for malignant glioma: long-term follow-up of a multicentre controlled trial. *Acta Neurochirurgica,* 148**,** 269-275.

WHITTLE, I. R. (2002) Surgery for gliomas. *Current Opinion in Neurology,* 15**,** 663-669.

WHITTLE, I. R. (2004) The dilemma of low grade glioma. *Journal of Neurology, Neurosurgery and Psychiatry,* 75 ii31-ii36.

WHITTLE, I. R., PRINGLE, A.-M. & TAYLOR, R. (1998) Effects of resective surgery for left-sided intracranial tumours on language function: a prospective study. *Lancet* 351**,** 1014-1018.

WHITTLE, I. R., SHORT, D. M., DEIGHTON, R. F. & MCCULLOCH, J. (2007) Proteomic analysis of gliomas. *British Journal of Neurosurgery,* 21**,** 576-582.

WHITTLE, I. R., SMITH, C., NAVOO, P. & COLLIE, D. (2004) Meningiomas. *The Lancet,* 363**,** 1535-1543.

WHYTE, J., CURRY, C. & HALE, D. (1985) Inspection time and intelligence in dyslexic children. *Journal of Child Psychology and Psychiatry,* 26**,** 423-428.

WIDAMAN, K. F. & CARLSON, J. S. (1989) Procedural effects on performance on the Hick paradigm: Bias in reaction time and movement time parameters. *Intelligence,* 13**,** 63-85.

WIEMANN, C. M., CHACKO, M. R., TUCKER, J. C., VELASQUEZ, M. M., SMITH, P. B., DICLEMENTE, R. J. & VON STERNBERG, K. (2005) Enhancing recruitment and retention of minority young women in community-based clinical research. *Journal of Pediatric and Adolescent Gynecology,* 18**,** 403-407.

YOUNG, A. H., SAHAKIAN, B. J. & ROBBINS, T. W. (1999) The effects of chronic administration of hydrocortisone on cognitive function in normal male volunteers. *Psychopharmacology,* 145**,** 260-266.

ZBINDEN, L., PANDEY, R., SCHWINDACK, C., DEARY, I. J. & WHITTLE, I. R. (2006) Information processing in patients with intracranial tumours: A preliminary evaluation of the clinical utility of inspection time. *British Journal of Neurosurgery,* 20**,** 203-208.

ZIGMOND, A. S. & SNAITH, R. P. (1983) The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica,* 67**,** 361-370.

ZIMPRICH, D. & MARTIN, M. (2002) Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and Ageing,* 17**,** 690-695.

The Role of Inspection Time in the Management of Patients with Intracranial Tumours

Principal Investigator: Prof. Ian R. Whittle

I am currently recruiting brain tumour patients for the above study and would really appreciate your help. I need to see patients prior to surgery for up to an hour and during this time they will complete a number of different tasks. I will then see them post-surgery for around 25 minutes. I am also recruiting healthy patients who are undergoing spinal surgery as controls. If you could contact me with information about any suitable patients, this would be a great help. I am based in ECNO and my contact details are below.

Tel: ext 33267

E-mail: jscotla2@staffmail.ed.ac.uk

We are also looking for healthy volunteers to take part in the study as controls. Please pass this information to anyone who may be willing to participate.

Thank you!

Appendix B. Patient Information Sheet

Patient Information Form
Assessment of Information Processing and the Effects of Therapy

You are being invited to take part in a research study that provides information about how the brain copes with new information. We wish to determine whether the brain is slowed down by the presence of a tumour and if treatments improve the speed at which the brain processes information. The tasks you complete will not influence your normal course of treatment in any way.

If you choose to participate in this study, you will be required to complete a number of thinking tasks. Some of the tasks involve remembering things, some involve working things out, and some look at how fast your brain is working on that day. Each task is short. One of the tasks is performed in front of a computer. You will be asked to make a simple decision about a pattern that appears on the screen. You will be given further, precise instructions before each task to ensure you understand completely what is involved.

You will be asked to complete these tasks before therapy (surgery). This should take about 1 hour. You will also be asked to complete a shorter session which will take up to 30 minutes a few days after therapy (surgery). If you have a brain tumour, you will also be asked to take part in a third session which will take place when you attend your outpatient appointment. The tests you complete will not influence your normal course of treatment in any way.

The information we gain from these results will help to determine whether certain therapies improve patient functioning.

All the information that is obtained is entirely confidential and stored in protected databanks. Your data will not be identifiable to any other person. As part of this research programme, we shall let your GP know that you have participated in the study.
Please feel free to ask questions if anything is not clear or if you would like more information.

If you require further information on these tests and why they are performed you or your family can contact Dr. Robin Grant, Consultant Neurologist, Western General Hospital (0131 537 2084). He is not involved in the study and so will be able to give you independent advice.

Yours sincerely
IAN R. WHITTLE MD PhD FRACS FRCSE (SN) FRCPE
Forbes Professor of Surgical Neurology

Participant Information Form

Assessment of Information Processing and the Effects of Therapy

Principal Investigator: Professor Ian R. Whittle, Department of Clinical Neurosciences, Western General Hospital, Edinburgh.

You are being invited to take part in a research study that provides information about how the brain copes with new information. We wish to determine whether the brain is slowed down by the presence of a tumour and if treatments improve the speed at which the brain processes information. You have been asked to take part in this study as part of a group of healthy volunteers. By comparing your results with those of a group of patients who have a brain tumour, we hope to learn more about this group of patients.

If you choose to participate in this study, you will be required to complete a number of thinking tasks. Some of the tasks involve remembering things, some involve working things out, and some look at how fast your brain is working on that day. Each task is short. One of the tasks is performed in front of a computer. You will be asked to make a simple decision about a pattern that appears on the screen. You will be given further, precise instructions before each task to ensure you understand completely what is involved.

Taking part in this study will involve two separate sessions. The first of these should take up to 1 hour. The second session will take place three to seven days after the first session and will be shorter, lasting around 25-30 minutes. These sessions can take place at any time convenient to you.

The information we gain from these results will help to determine whether certain therapies improve patient functioning.

All the information that is obtained is entirely confidential and stored in protected databanks. Your data will not be identifiable to any other person.

Please feel free to ask questions if anything is not clear or if you would like more information.

Yours sincerely
IAN R. WHITTLE MD PhD FRACS FRCSE (SN) FRCPE
Forbes Professor of Surgical Neurology
Consultant Neurosurgeon

Appendix D. Letter to recruitment panel

Edinburgh Centre of Neuro-Oncology
Western General Hospital
Crewe Road South
Edinburgh
EH4 2XU
Date

Dear ,

As a registered member of the recruitment panel for the Department of Psychology, University of Edinburgh, we are writing to ask you to participate as a healthy volunteer in a study currently being carried out at the Western General Hospital, Edinburgh.

The study is being carried out in order to examine how patients with a brain tumour perform on a number of different thinking tasks carried out on two separate occasions - before and after surgery to remove the tumour. Your participation will allow us to compare the group of patients with a brain tumour with a group of healthy volunteers.

Participating in this study will involve two separate testing sessions both of which will take place at the hospital:
The first session will last up to an hour and during this time, you will complete a number of short thinking tasks.

The second session will take place 3-7 days later and will last around 20-30 minutes. Here you will be given some of the tasks you completed during session 1.

Both sessions will take place at a pre-arranged time convenient to you and can take place during the day, in the evening or at the weekend depending on your preference. We can reimburse your travel to and from the hospital.

We are looking for people who speak English as a first language and who have never had any kind of brain disorder.

Your participation in this study would be greatly appreciated. If you are willing to take part or if you require any further information about what taking part would involve please email me at jscotla2@staffmail.ed.ac.uk or telephone me on 0131 537 3267.

I look forward to hearing from you.

Yours sincerely

Jennifer Scotland
Research Associate

Patient Consent Form
Assessment of Information Processing

Principal Investigator: Professor Ian R. Whittle, Department of Clinical Neurosciences, Western General Hospital, Edinburgh.

- I have read the information sheet that has been provided to me, and this Consent Form, and have been given the opportunity to ask questions about them. I am satisfied that I have all the information that I need to provide informed consent.

- I understand that my doctor will be informed of my participation in this study and know that he/she will be provided with a routine clinical report.

- I know that I am under no obligation to take part in this study and can withdraw at any time without my care being compromised.

- I agree that any test results obtained during assessment may be stored and processed using computers and after the study is complete they may be copied onto a permanent record and may be studied again at a later time.

- I agree that information gathered during this study may be shared with other medical and scientific researchers, subject to laws and University policies intended to safeguard my privacy.

- I agree to participate in this study.

Signature of patient:

Name of patient (please print in block capitals):

Witnessed by:

Name of witness (please print in block capitals):

Date:

Consent Form
Assessment of Information Processing

Principal Investigator: Professor Ian R. Whittle, Department of Clinical Neurosciences, Western General Hospital, Edinburgh.

- I have read the information sheet that has been provided to me, and this Consent Form, and have been given the opportunity to ask questions about them. I am satisfied that I have all the information that I need to provide informed consent.

- I know that I am under no obligation to take part in this study and can withdraw at any time without giving a reason for doing so.

- I agree that any test results obtained during assessment may be stored and processed using computers and after the study is complete they may be copied onto a permanent record and may be studied again at a later time.

- I agree that information gathered during this study may be shared with other medical and scientific researchers, subject to laws and University policies intended to safeguard my privacy.

- I agree to participate in this study.

Signature of participant:

Name of participant (please print in block capitals):

Witnessed by:

Name of witness (please print in block capitals):

Date:

ORDER OF TEST ADMINISTRATION
Session 1 -Baseline

No.	Test
1	Rey Auditory-Verbal Learning Test (I – VI)
2	Inspection Time*
3	Rey Auditory-Verbal Learning Test (VII)
4	National Adult Reading Test (NART)
5	Trail Making Test Part B
6	Verbal Fluency
7	Hospital Anxiety and Depression Scale
8	Letter-Number Sequencing
9	Quality of Life Questionnaire
10	EFIT Williams Delayed Recall Test (Part 1)*
11	EFIT Nine Hole Peg Test*
12	Digit Symbol-Coding*
13	EFIT Williams Delayed Recall Test (Part 2)*
14	EFIT Boston Aphasia Severity Rating Scale*
15	Barthel Disability Index
16	Karnofsky Performance Scale
17	EFIT Timed Ten Metre Walk*

* denotes tests completed at session 2.

Appendix H. Rey Auditory Verbal Learning Test Word Lists

List A (Learning)	List B (Interference)
DRUM	DESK
CURTAIN	RANGER
BELL	BIRD
COFFEE	SHOE
SCHOOL	STOVE
PARENT	MOUNTAIN
MOON	GLASSES
GARDEN	TOWEL
HAT	CLOUD
FARMER	BOAT
NOSE	LAMB
TURKEY	GUN
COLOUR	PENCIL
HOUSE	CHURCH
RIVER	FISH

CHORD

ACHE

DEPOT

AISLE

BOUQUET

PSALM

CAPON

DENY

NAUSEA

DEBT

COURTEOUS

RAREFY

EQUIVOCAL

NAÏVE

CATACOMB

GAOLED

THYME

HEIR

RADIX

ASSIGNATE

HIATUS

SUBTLE

PROCREATE

GIST

GOUGE

SUPERFLUOUS

SIMILE

BANAL

QUADRUPED

CELLIST

FAÇADE

ZEALOT

DRACHM

AEON

PLACEBO

ABSTEMIOUS

DÉTENTE

IDYLL

PUERPERAL

AVER

GAUCHE

TOPIARY

LEVIATHAN

BEATIFY

PRELATE

SIDEREAL

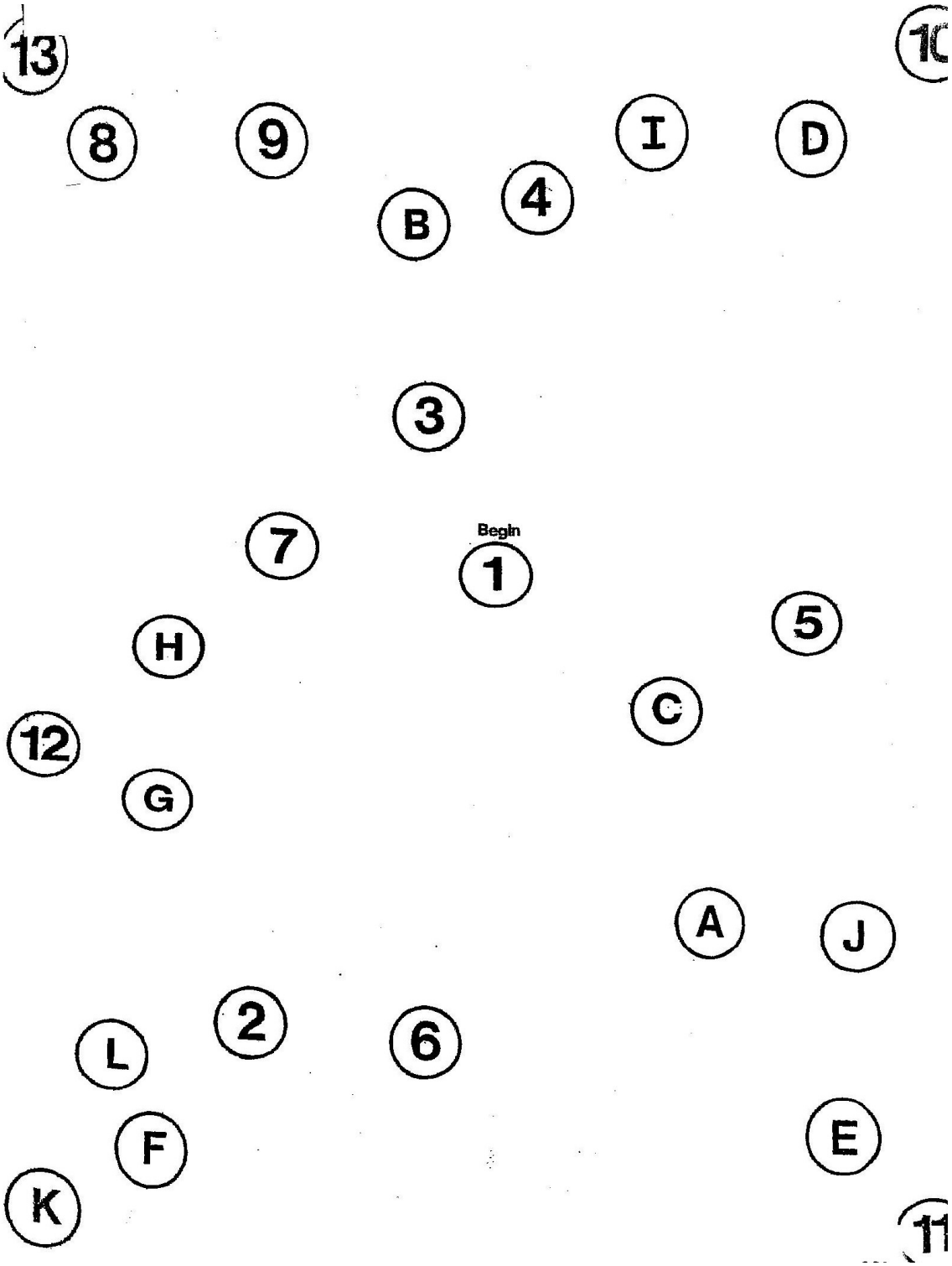
DEMESNE

SYNCOPE

LABILE

CAMPANILE

Appendix J. Trail Making Test Part B Test Sheet



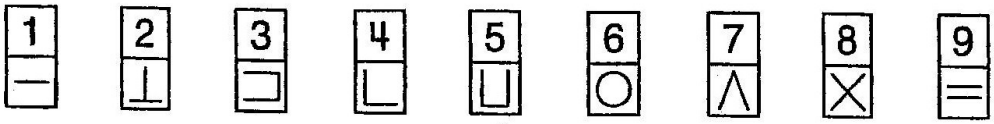
Appendix K. Verbal Fluency Test – Response Recording Sheet

C	F	L
SCORE:	**SCORE:**	**SCORE:**

TOTAL: _____

Appendix L. Digit Symbol Coding Test Sheet

Digit Symbol—Coding



Sample Items

2	1	3	7	2	4	8	2	1	3	2	1	4	2	3	5	2	3	1	4

5	6	3	1	4	1	5	4	2	7	6	3	5	7	2	8	5	4	6	3

7	2	8	1	9	5	8	4	7	3	6	2	5	1	9	2	8	3	7	4

6	5	9	4	8	3	7	2	6	1	5	4	6	3	7	9	2	8	1	7

9	4	6	8	5	9	7	1	8	5	2	9	4	8	6	3	7	9	8	6

2	7	3	6	5	1	9	8	4	5	7	3	1	4	8	7	9	1	4	5

7	1	8	2	9	3	6	7	2	8	5	2	3	1	4	8	4	2	7	6

Appendix M. Letter-Number Sequencing Test Sheet

	Trial	Item	Response	TrialScore (0 or 1)	Item Score (0, 1, 2,or 3)
1	1	L – 2	2 – L		
	2	**6–P**	6 – P		
	3	**B – 5**	5 – B		
2	1	**F – 7 – L**	7 – F – L		
	2	**R – 4 – D**	4 – D – R		
	3	**H – 1 – 8**	1 – 8 – H		
3	1	**T – 9 – A – 3**	3 – 9 – A – T		
	2	**V – 1 – J – 5**	1 – 5 – J- V		
	3	**7 – N – 4 – L**	4 – 7 – L – N		
4	1	**8 – D – 6 – G – 1**	1 – 6 - 8 – D – G		
	2	**K – 2 – C – 7 – S**	2 – 7 – C – K – S		
	3	**5 – P – 3 – Y – 9**	3 – 5 – 9 – P –Y		
5	1	**M – 4 – E – 7 – Q – 2**	2 – 4 – 7 – E – M – Q		
	2	**W – 8 – H – 5 – F – 3**	3 – 5 – 8 – F – H – W		
	3	**6 – G – 9 – A - 2 – S**	2 – 6 – 9 – A – G – S		
6	1	**R – 3 – B - 4- Z – 1– C**	1 – 3 – 4 – B – C – R – Z		
	2	**5 – T – 9 – J – 2 – X- 7**	2 – 5 – 7 – 9 - J – T – X		
	3	**E – 1 – H –8 – R– 4- D**	1 – 4 – 8 – D – E – H – R		
7	1	**5 – H –9- S- 2- N- 6–A**	2 – 5 –6 – 9 –A –H –N –S		
	2	**D –1 –R –9- B- 4- K- 3**	1 –3– 4 –9 –B –D –K – R		
	3	**7 –M –2 –T -6- F- 1- Z**	1 –2 –6 –7 – F –M –T – Z		
TOTAL RAW SCORE (MAX. = 21)					

HAD Scale

This questionnaire is designed to help us know how you feel. Read each item and place a firm tick in the box opposite the one that comes closest to how you have been feeling **in the past week**.

Try not to take too long over your replies; your immediate reaction to each item will probably be more accurate than a long thought out response.

Tick only one box in each section.

I feel tense or 'wound up':	**TICK**
Most of the time	
A lot of the time	
Some of the time, occasionally	
Not at all	

I feel as if I am slowed down:	**TICK**
Nearly all the time	
Very often	
Sometimes	
Not at all	

I still enjoy the things I used to enjoy:	**TICK**
Definitely as much	
Not quite as much	
Only a little	
Hardly at all	

I get a sort of frightened feeling like 'butterflies' in the stomach:	**TICK**
Not at all	
Occasionally	
Quite often	
Very often	

I get a sort of frightened feeling as if something awful is about to happen:	**TICK**
Very definitely and quite badly	
Yes, but not too badly	
A little, but it doesn't worry me	
Not at all	

I have lost interest in my appearance:	**TICK**
Definitely	
I don't take so much care as I should	
I may not take quite as much care	
I take just as much care as ever	

I can laugh and see the funny side of things:	**TICK**
As much as I always could	
Not quite so much now	
Definitely not so much now	
Not at all	

I feel restless as if I have to be on the move:	**TICK**
Very much indeed	
Quite a lot	
Not very much	
Not at all	

Worrying thoughts go through my mind:	**TICK**
A great deal of the time	
A lot of the time	

I look forward with enjoyment to things:	**TICK**
As much as I ever did	
Rather less than I used to	

From time to time, but not too often	
Only occasionally	

Definitely less than I used to	
Hardly at all	

I feel cheerful:	**TICK**
Not at all	
Not often	
Sometimes	
Most of the time	

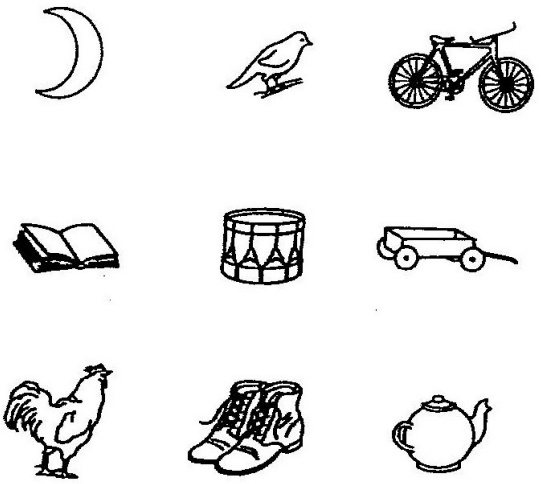
I get sudden feelings of panic:	**TICK**
Very often indeed	
Quite often	
Not very often	
Not at all	

I can sit at ease and feel relaxed:	**TICK**
Definitely	
Usually	
Not often	
Not at all	

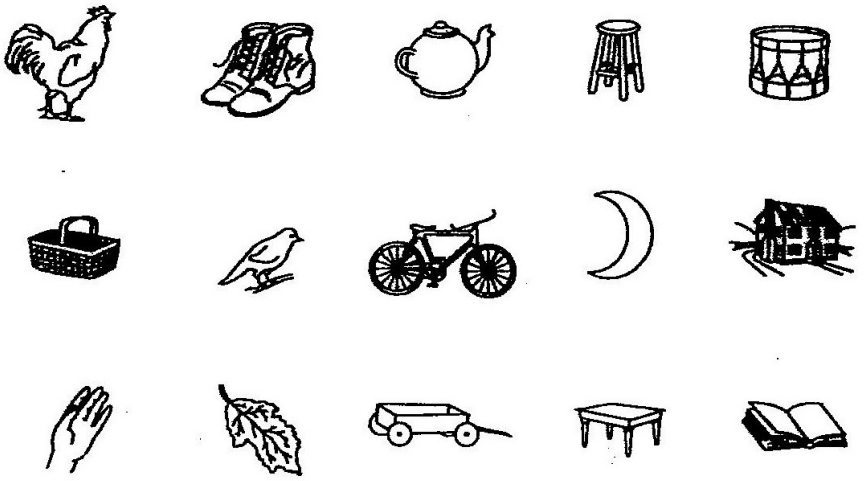
I can enjoy a good book or radio or TV programme:	**TICK**
Often	
Sometimes	
Not often	
Very seldom	

A count	
D count	
Total	

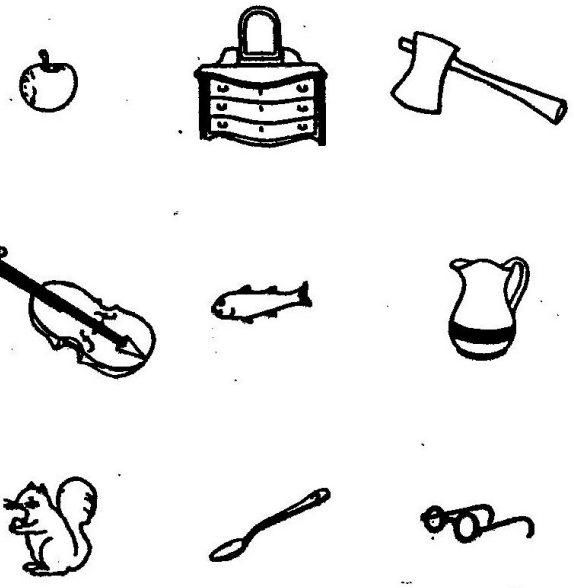
A



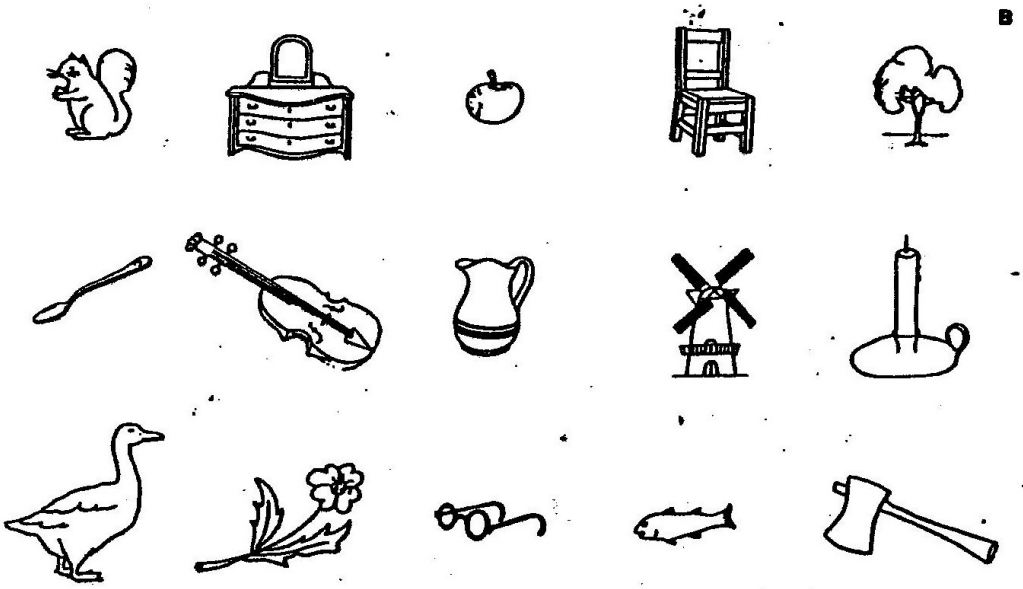
A

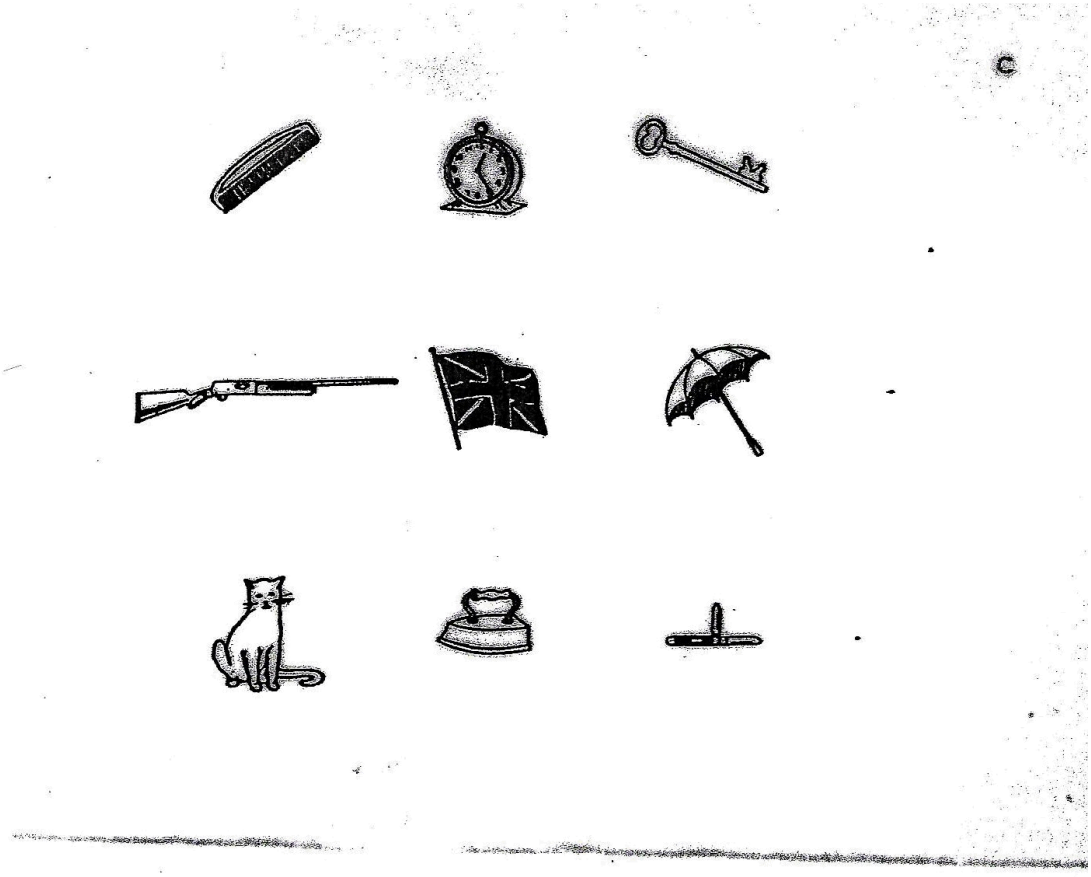


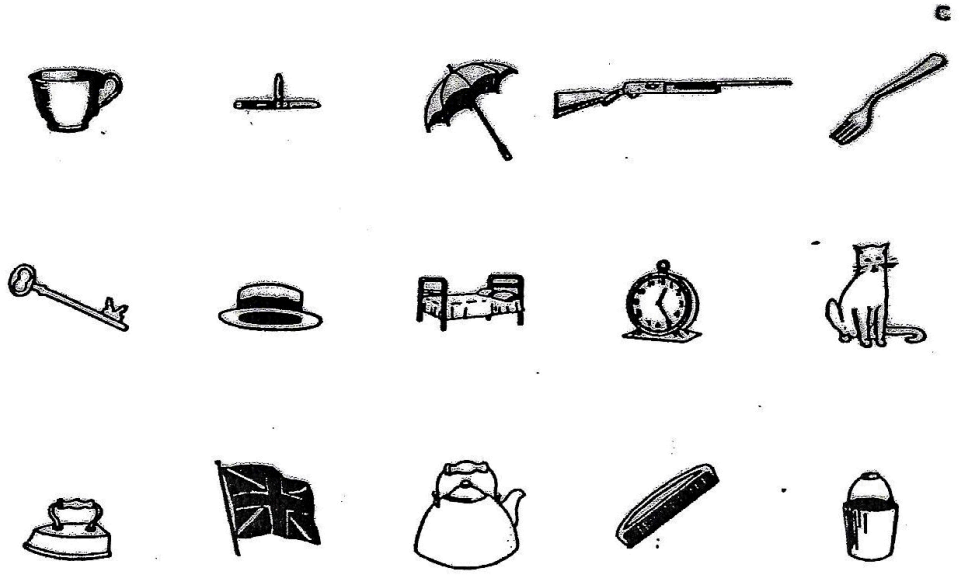
B



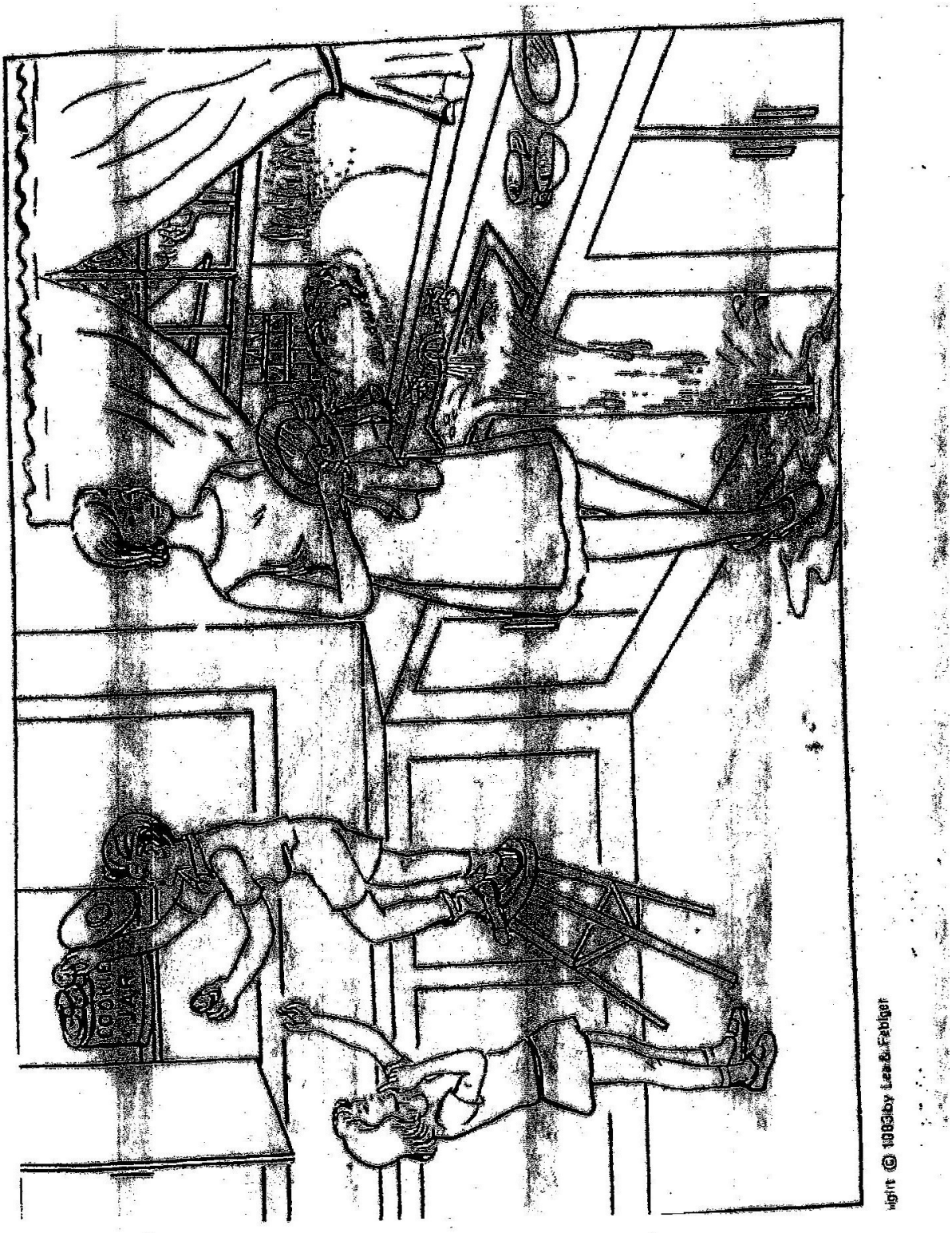
B







Appendix P. Boston Aphasia Severity Rating Scale ‘Cookie Theft’ Picture



EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in **your name:** _____

Your birth date (Day, Month, Year): _____

Today's date (Day, Month, Year): _____

Did you fill in the questionnaire yourself? No ☐ Yes ☐

	No	**Yes**
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2
2. Do you have any trouble taking a long walk?	1	2
3. Do you have any trouble taking a short walk outside of the house?	1	2
4. Do you need to stay in a bed or a chair during the day?	1	2
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2

During the past week:	Not at All	A Little	Quite a Bit	Very Much
6. Were you limited in doing either your work or other daily activities?	1	2	3	4
7. Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8. Were you short of breath?	1	2	3	4
9. Have you had pain?	1	2	3	4
10. Did you need to rest?	1	2	3	4
11. Have you had trouble sleeping?	1	2	3	4
12. Have you felt weak?	1	2	3	4
13. Have you lacked appetite?	1	2	3	4
14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
16. Have you been constipated?	1	2	3	4
17. Have you had diarrhoea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you:

29. How would you rate your overall health during the past week?	1 2 3 4 5 6 7 Very Poor Excellent
30. How would you rate your overall quality of life during the past week?	1 2 3 4 5 6 7 Very Poor Excellent

During the past week:	**Not at All**	**A Little**	**Quite a Bit**	**Very Much**
31. Did you feel uncertain about the future?	1	2	3	4
32. Did you feel you had setbacks in your condition?	1	2	3	4
33. Were you concerned about disruption of family life?	1	2	3	4
34. Did you have headaches?	1	2	3	4
35. Did your outlook on the future worsen?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
36. Did you have double vision?	1	2	3	4
37. Was your vision blurred?	1	2	3	4
38. Did you have difficulty reading because of your vision?	1	2	3	4
39. Did you have seizures?	1	2	3	4
40. Did you have weakness on one side of your body?	1	2	3	4
41. Did you have trouble finding the right words to express yourself?	1	2	3	4
42. Did you have difficulty speaking?	1	2	3	4
43. Did you have trouble communicating your thoughts?	1	2	3	4
44. Did you feel drowsy during the daytime?	1	2	3	4
45. Did you have trouble with your co-ordination?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
46. Did hair loss bother you?	1	2	3	4
47. Did itching of your skin bother you?	1	2	3	4
48. Did you have weakness of both legs?	1	2	3	4
49. Did you feel unsteady on your feet?	1	2	3	4
50. Did you have trouble in controlling your bladder?	1	2	3	4

EUROQOL

To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked '100' and the worst state you can imagine is marked by '0'.

We would like you to indicate on this scale how good or bad your health is today, in your opinion. Please do this by ticking a box on the scale indicating how good or bad your current health state is.

Your Current Health State:

	100
	90
	80
	70
	60
	50
	40
	30
	20
	10
	0

BARTHEL DISABILITY INDEX

Feeding

	2 = Independent	Able to use any necessary device: eats in a reasonable time: able to butter on his/her own.
	1 = Needs help	Cutting or spreading butter.
	0 = Dependent	Needs to be fed.

Bathing

	1 = Independent	Able to wash self all over; maybe by using a shower, full bath or stable getting in or out of shower room.
	0 = Dependent	Needs some help.

Dressing

	2 = Independent	Includes tying shoelaces, 7 buttons.
	1 = Needs help	Able to dress by self but needs help tying shoes, 7 fasteners.
	0 = Dependent	Needs to be dressed.

Grooming

	1 = Independent	Washing hands/face, combing hair, brushing teeth.
	0 = Dependent	Needs some help.

Mobility

	3 = Independent	Can walk 50m. May use aid except roller, speed not important.
	2 = Needs help	Needs help of one person, including help standing.
	1 = In wheelchair	Independent.
	0 = Immobile	Including being wheeled by another.

Stairs

	2 = Independent	Must carry walking aid if used; needs no help.
	1 = Needs help	Manages with help. Physical or verbal supervision, carrying aid.
	0 = Dependent	Unable. Cannot negotiate stairs, needs lift.

Toilet

	2 = Independent	Able to handle clothes, wipe self, flush toilet, empty commode completely.
	1 = Needs help	Able to manage with minor help balancing, handling clothes or toilet.
	0 = Dependent	Needs major assistance.

Bed Chair

	3 = Independent	Needs no help in moving from bed to (wheel)chair and vice versa. In wheelchair if necessary.
	2 = Minimal help	Verbal supervision, minor physical help from spouse.
	1 = Major help	Able to sit unaided but needs much help of one/two persons.
	0 = Dependent	Unable to sit. Needs hoist or lift by two persons.

Bowels

	2 = No accidents.
	1 = Occasional accidents/help with enema.
	0 = Incontinent.

Bladder

	2 = No accidents.
	1 = Occasional incontinence or device used.
	0 = Incontinent.

Barthel Score	

KARNOFSKY PERFORMANCE SCALE

Able to carry on normal activity. No special care needed.		100	Normal, no complaints or evidence of disease.
		90	Normal activity, minor signs and symptoms.
		80	Normal activity with effort, signs and symptoms of disease.
Unable to work. Able to live at home and care for most personal needs; a varying amount of assistance needed.		70	Cares for self: unable to do normal activity or work.
		60	Requires occasional assistance, able to care for most needs.
		50	Requires considerable assistance and frequent medical care.
Unable to care for self; requires equivalent of institutional or hospital care. Disease may be progressing rapidly.		40	Disabled; requires special medical care and assistance.
		30	Severely disabled. Hospitalisation indicated.
		20	Very sick. Hospitalisation necessary. Active support Rx necessary.
		10	Moribund.
		0	Dead.

Acta Neurochir
DOI 10.1007/s00701-009-0371-9

CLINICAL ARTICLE

Recruitment difficulties in brain tumour patients cause participation bias: findings from a neuropsychological study of adult inpatients with supratentorial intracranial tumours

Jennifer L. Scotland · Rustam Al-Shahi Salman · Ian J. Deary · Ian R. Whittle

Received: 2 March 2009 / Accepted: 10 April 2009

Abstract
Purpose Patients who participate in questionnaire surveys, clinical studies and clinical trials can be different from patients who do not participate. The occurrence and direction of this response, participation or ascertainment bias is unpredictable, and can harm the external validity of medical research.
Methods We compared the characteristics of patients with intracranial tumours who participated in a psychological study of inspection time with the characteristics of patients who did not participate for a number of reasons.
Results Of 178 newly diagnosed adults with intracranial tumours, 136 (76%) were eligible, of whom 76 (56%) participated and 34 (25%) declined. There were no significant differences in terms of age and sex of the patients who participated and those who declined. When the participation group was combined with those who were ineligible and those who declined, the majority of patients in the combined cohort (n=152) had a WHO grade III or IV glioma (high-grade glioma) (48.0%), and only 13.2% had a WHO grade I or II glioma (low-grade glioma). However, only 38.2% of those who participated had a WHO grade III or IV glioma, and 23.7% had a WHO grade I or II glioma. Comparisons of the participation vs. ineligible and declined groups revealed there was a significant difference (p=0.002) between the ratio of high-grade to low-grade gliomas in the total and recruited cohorts. Comparisons of only the participation vs. declined groups approached significance (p=0.051). WHO grade III and IV glioma patients were under-represented, and WHO grade I or II glioma patients were over-represented in the study group.
Conclusions Noninterventional, non-therapeutic applied neuropsychological studies in neuro-oncology are susceptible to bias since the spectrum of neuropathologies in recruited patients can be significantly different from that of the total cohort. These data could help anticipate recruitment rates for applied neuropsychological studies in clinical neuro-oncology and may help anticipate likely selection biases amongst those who participate.

Keywords Glioma · Neuropsychology · Recruitment bias

Introduction

Recruiting participants into clinical research studies is often a difficult task and can ultimately result in an unsuccessful clinical trial [24]. The causes of recruitment difficulties include *inter alia* the need to recruit vulnerable or seriously ill patients, the necessity to adhere to strict ethical guidelines, indirect pathways to patient recruitment and changes in staff involved in trial recruitment. A well-planned recruitment strategy involving collaboration with other health-care professionals and flexibility is essential [14].

This paper was presented in part at The British Neurosurgical Research Group Meeting in Manchester, February 2007, and the British Neuro-oncological Society meeting in Preston, June 2008.

J. L. Scotland (✉) · R. Al-Shahi Salman · I. R. Whittle
Division of Clinical Neurosciences, University of Edinburgh, Edinburgh, UK
e-mail: jscotla2@staffmail.ed.ac.uk

I. J. Deary
Division of Psychology and Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

Published online: 14 May 2009

Selective recruitment, in addition to other factors, can compromise the integrity of a clinical study, and low participation rates from patients who *are* approached increases the likelihood of a sample bias [1, 23]. Thus, in a prospective, observational study of brain arteriovenous malformations, there was a significant difference between those who consented and those who either refused or could not consent in terms of some clinical features, likelihoods of treatment and likelihood of intracranial haemorrhage [1].

Here we describe our experience in recruiting neurosurgical inpatients into an ethically approved, prospective, externally funded, non-therapeutic, non-interventional, applied neuropsychological study in brain tumour patients. Based on the results of a pilot study [25], this larger prospective study was designed and powered to evaluate the potential benefits of measuring visual inspection time in neuro-oncological patients. The recruitment timetable had been constructed based on an estimated 90-100 supratentorial brain tumour patients per year, with a 70% recruitment rate over 30 months. Estimates of recruitment numbers and rates were based on several previous prospective neuro-oncological studies in the department, one an interventional therapeutic study with 100% compliance [19], two non-interventional neuroradiological studies with 100% compliance [2, 12] and another non-interventional neuropathological study with >96% compliance [21], and a prospective audit that suggested about 25% of patients would have cognitive or other problems that rendered them ineligible for neuropsychological studies [6]. A recruitment log of all patients approached was kept from the onset.

After it became apparent that recruitment was falling behind schedule, a dedicated study was performed to analyse the extent of the failure of planned recruitment and to evaluate whether recruitment difficulties may cause bias in the sample.

Methods

The study was focused on patients with newly diagnosed supratentorial intracranial tumours who were to have either a biopsy or surgical resection, and control patients having elective spinal surgery for neurodegenerative conditions. Participation in the study, which was ethically approved (LREC/2002/5/4), involved the completion of: (1) a battery of neuropsychological tests before surgery, including inspection time (total duration 40-60 min), and (2) a follow-up testing session postoperatively, prior to discharge, in which the patient repeats a number of the tasks (total duration 30 min). A third session for only the brain tumour patients took place when the patient returned for a follow-up outpatient appointment (duration 30 min).

In order to optimise recruitment, the research associate (JLS) consulted the admissions diary and the pre-admission staff for relevant wards on a daily basis in order to obtain names of suitable elective cases. Recruitment posters detailing information about the study were placed on the wards. All ward medical and paramedical staff were informed about the study, and were also asked directly by IRW and by the research associate to inform them of any suitable admissions. A recruitment log was kept, tracking all potential brain tumour participants, their demographic characteristics and the reasons for their non-participation.

Data were analysed using 2×2 tables and nonparametric statistics on SPSS software.

Results

A total of 178 newly diagnosed patients were admitted for surgical biopsy or resection of a supratentorial intracranial tumour to the Department of Clinical Neurosciences between April 2006 and December 2007. Of these, 136 (76%) fulfiled study eligibility criteria. Forty-two (24.0%) of the 178 patients were ineligible to take part in the study for medical reasons (severe visual acuity or visual field defects, dysphasia or impaired mental status that precluded obtaining informed consent, or sensorimotor impairments that prevented the patient from being transported to the testing office), or because of co-morbidities such as Down syndrome, schizophrenia or chronic alcoholism. Of the cohort eligible for the study, only 56% participated, since 25% (34 patients) declined and 19% (26 patients) were 'missed' because the research associate was on annual leave (n=13), there was a communication failure between ward medical and nursing staff that left no time for testing (n=7), or the patients were either admitted or transferred to the ward very late on the day before their operation (n=6) (Fig. 1).

The median age of those brain tumour patients who participated was 52 years (range 17–80), which was not significantly different from those who declined (54 years; range 31–76). Of those brain tumour patients who declined to go in the study, 23 (68%) of the 34 did not give any specific reason for refusal. The remaining 11 patients (32%) stated anxiety about their impending surgery as the reason for declining. There was no significant association between the patient gender and whether or not they agreed to participate in the study (χ^2 (1)=3.11, p=0.78).

Of the total cohort who either participated, declined or were ineligible (n=152), the majority were subsequently found to have a high-grade glioma (48.0%, WHO III-IV), low-grade glioma (13.2%; WHO I or II), meningi-

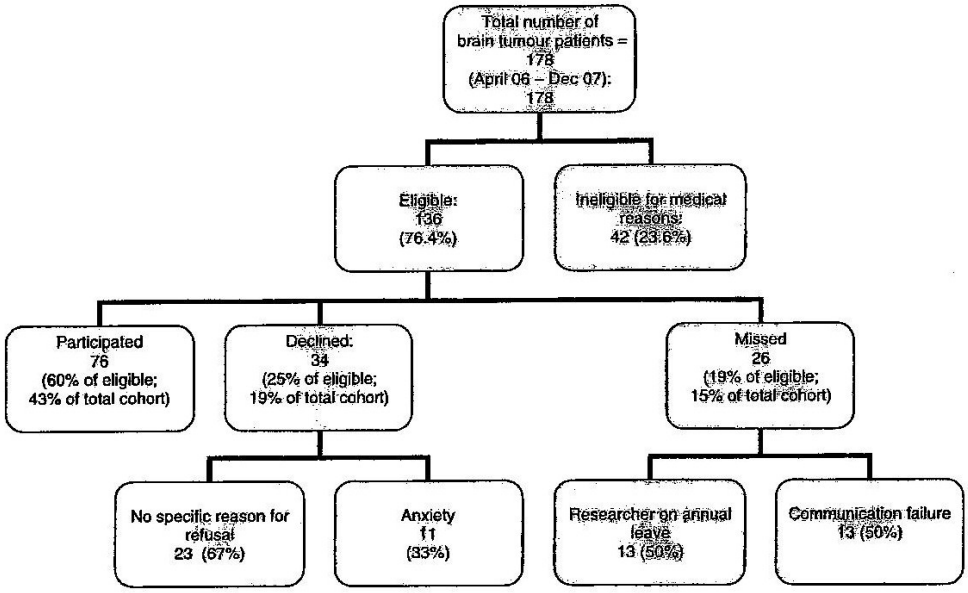


Fig. 1 Flowchart detailing brain tumour patient cohort, eligibility for the study and recruitment

oma (17.1%) or metastasis (9.9%) (see Table 1). However, in the study cohort only 38.2% had a WHO grade III or IV glioma, and 24% had a WHO grade I or II low-grade glioma. This anomaly arose because 66.7% of the patients in the medically 'excluded' ineligible group had WHO grade III or IV gliomas, and only 2.4% of ineligible patients were found to have a WHO grade I or II gliomas.

The 'declined' and 'ineligible' groups were combined to form a single 'non-participation' group and were compared in terms of tumour type with the participating group. This analysis revealed a significant association between the group and tumour type: $\chi^2(4)=16.79$, $p=0.002$. This confirms that more WHO grade III or IV glioma patients than expected were not entered into the study and, conversely, a higher number of WHO grade I or II glioma patients than expected did participate.

When the group who declined to take part was compared with the participation group, the association between group and tumour type approached the conventional level of significance [X^2 (4)=9.46; p=0.051]. Again, this highlights the differences between those who did and did not agree to participate in the study.

Discussion

The patient recruitment phase is often the most problematic aspect of carrying out a research project [9]. In this applied

Table 1 Number of patients with each tumour type in the total cohort (not including missed patients) and subgroups of those that participated, declined or were ineligible

	Tumour type					
		High-grade glioma (WHO III-IV)	Low-grade glioma (WHO I or II)	Meningioma	Metastasis	Other/not known
Group	Total n (% within group)	73 (48.0%)	20 (13.2%)	26 (17.1%)	15 (9.9%)	18 (11.8%)
	Participated n (% within group)	29 (38.2%)	18 (23.7%)	11 (14.5%)	8 (10.5%)	10 (13.2%)
	Declined n (% within group)	16 (47.1%)	1 (2.9%)	10 (29.4%)	4 (11.8%)	3 (8.8%)
	Ineligible n (% within group)	28 (66.7%)	1 (2.4%)	5 (11.9%)	3 (7.1%)	5 (11.9%)

The study subgroup was substantially different to the non-participators (declined and ineligible, p=0.002). High-grade glioma included WHO III and IV tumours, low-grade glioma WHO II tumours, and other/not known included a range of tumours and a small number of patients in whom surgery was subsequently not performed

neuropsychological project in patients with supratentorial brain tumours, the decline rate was higher in the brain tumour group when compared to a recent interventional brain tumour clinical trial [17], two non-interventional, non-testing (demanding) clinical studies [2, 12] and an ongoing non-interventional non-'testing' brain tumour clinicopathological study [21] that were also conducted in the unit. Where a reason for declining to participate was given, which only occurred in 32% of cases, the anxiety associated with a new diagnosis of brain tumour combined with the anxiety about having to perform a range of neuropsychological tests was mentioned. Certainly the problems of performing neuropsychological tests in brain tumour patients have been well documented qualitatively but not quantitatively [15]. Potential problems with neuropsychological evaluation in brain tumour patients would seem maximal in WHO grade III and IV glioma patients before the first surgery since most studies have been done successfully on both low- and high-grade glioma patients during follow-up [3, 5, 7, 11].

However, the non-therapeutic nature of the study may also have increased decline rates, since overall only 56% of all eligible brain tumour patients (43% of the entire cohort) were recruited into the study. This is comparable with data from a community-based clinical research study involving young women of an ethnic minority in which 47% of all eligible participants were successfully recruited [22]. Similarly, a mean enrolment proportion of 32.7% across 112 different recruitment sites into a Cardiac Arrhythmia Suppression Trial (CAST) has also been reported [13]. In the recent ISAT trial only 2,143 (22.42%) cases out of 9,559 patients with intracranial aneurysms were recruited into this randomised interventional study [8]. Generally, however, there are very few data in the literature about what influences participation in various clinical trials.

Analysis of the proportion of patients with each tumour type (WHO grade III and IV glioma, WHO grade I and II glioma, meningioma, metastasis or other tumour) in each of the recruitment groups (participated, declined, excluded) revealed significantly more patients with a WHO grade III or IV glioma diagnosis in the ineligible and declined groups compared with the proportion of WHO III or IV glioma patients who participated. This finding is not altogether unexpected given the tendency for patients with malignant gliomas to have more focal and global brain dysfunction than patients with lower grade gliomas or meningiomas [18]. Additionally, these patients tend to present more frequently on an urgent and emergency basis, and as such are often admitted to the hospital ward for only a short while before surgical intervention. Conversely, most patients with WHO grade I and II glioma were recruited into the study. Since most WHO grade I and II glioma patients are treated surgically on an elective basis, and tend to have no focal neurological deficit or significant cognitive dysfunction, this is not surprising [20]. However, as a result of all these factors, WHO grade III and IV glioma patients are under-represented, and WHO grade I and II glioma patients are over-represented. Indeed, it has been previously noted in a prospective neurosurgical study that those recruited were significantly less likely to be either dead or disabled than those not recruited [1]. Another study evaluating stroke also found that those consenting to the study had significantly lower inpatient mortality than those not recruited [16]. Such a disparity may well bias the final outcome of our study, which was to determine differences between inspection time in patients with an intracranial tumour prior to surgery and two control groups. Given the high likelihood that patients with either WHO III or IV gliomas are likely to perform less well than patients with lower grade gliomas [3, 4], the tumour cohort recruited will score better than a truly representative cohort. Further bias in this study has also occurred because nine of the patients with glioblastoma recruited had their inspection time scores removed since they did not score in the minimally required range.

Although previous studies of cognition in WHO grade III and IV glioma patients generally report greater impairment than those studies focused on WHO grade I and II glioma patients [3, 4], these studies may actually underestimate the cognitive effects of brain tumours in WHO grade II and IV patients. This will occur since many patients, both in this and other studies, with more severe symptoms decline or are unable to participate [1, 10]. Our experiences highlight the methodological difficulties and potential consent or participation bias when recruiting patients into a non-interventional, neuropsychological study in neuro-oncology.

Acknowledgements This study received funding from the Chief Scientist Office (CSO) (CZH/4/232).

References

1. Al-Shahi R, Vousden C, Warlow C, Scottish Intracranial Vascular Malformation Study (SIVMS) Steering Committee (2005) Bias from requiring explicit consent from all participants in observational research: prospective, population based study. BMJ 331:942. doi:10.1136/bmj.38624.397569.68
2. Bastin ME, Carpenter TK, Armitage A, Sinha S, Wardlaw JM, Whittle IR (2006) Effects of dexamethasone on cerebral perfusion and water diffusion in patients with high-grade glioma. AJNR Am J Neuroradiol 27(2):402–408
3. Bosma I, Vos MJ, Heimans JJ, Taphoorn MJB, Aaronson NK, Postma TJ, van der Ploeg HM, Muller M, Vendertop WP, Slotma BJ, Klein M (2007) The course of neurocognitive functioning in high-grade glioma patients. Neuro-oncol 9(1):53–62. doi:10.1215/15228517-2006-012

4. Correa DD, Shi W, Thaler HT, Cheung AM, DeAngelis LM, Abrey LE (2008) Longitudinal cognitive follow-up in low grade gliomas. J Neurooncol 86:321–327. doi:10.1007/s11060-007-9474-4
5. Ek L, Smits A, Pahlson A, Almvist O (2005) Analysis of cognitive dysfunction in patients with low-grade glioma. J Clin Psychol Med Settings 12(2):165–173. doi:10.1007/s10880-005-3276-7
6. Grant R (2004) Overview: Brain tumour diagnosis and management. Royal college of Physician guidelines. J Neurol Neurosurg Psychiatry 75:ii18–ii23 (suppl 2)
7. Hahn CA, Dunn RH, Logue PE, King JH, Edwards CL, Halperin EC (2003) Prospective study of neuropsychologic testing and quality-of-life assessment of adults with primary malignant brain tumors. Int J Radiat Oncol Biol Phys 55(4):992–999. doi:10.1016/S0360-3016(02)04205-0
8. International Subarchnoid Aneurysm Trial (ISAT) Collaborative Group (2002) International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomized trial. Lancet 360:1267–1274. doi:10.1016/S0140-6736(02)11314-6
9. Isaacman DJ, Reynolds EA (1996) Effect of a research nurse on patient enrolment in a clinical study. Pediatr Emerg Care 12 (5):340–342. doi:10.1097/00006565-199610000-00004
10. Jackson C, Crossland L, Dennis M, Wardlaw J, Sudlow C (2008) Assessing the impact of the requirement for explicit consent in a hospital-based stroke study. Q J Med 101:281–289.
11. Klein M, Taphoorn MJB, Hiemans JJ, van der Ploeg HM, Vandertop WP, Smit EF, Leenstra S, Tulleken CAF, Boogerd W, Belderbos JSA, Cleijne W, Aaronson NK (2001) Neurobehavioural status and health-related quality of life in newly diagnosed high-grade glioma patients. J Clin Oncol 19(20):4037–4047
12. Schwindack C, Siminotto E, Meyer M, McNamara A, Marshall I, Wardlaw JM, Whittle IR (2005) Real-time functional magnetic resonance imaging (rt-fMRI) in patients with brain tumours: preliminary findings using motor and language paradigms. Br J Neurosurg 19(1):25–32. doi:10.1080/02688690500089621
13. Shea S, Bigger JT Jr, Campion J, Fleiss JL, Rolnitzky LM, Schron E, Gorkin L, Handshaw K, Kinney MR, Branyon M (1992) Enrollment in clinical trials: institutional factors affecting enrolment in the Cardiac Arrhythmia Suppression Trial (CAST). Control Clin Trials 13:466–486. doi:10.1016/0197-2456(92)90204-D
14. Sullivan-Bolyai S, Bova C, Deatrick JA, Knafl K, Grey M, Leung K, Trudeau A (2007) Barriers and strategies for recruiting study participants in clinical settings. West J Nurs Res 29(4):486–500. doi:10.1177/0193945907299658
15. Taphoorn MJB, Klein M (2004) Cognitive deficits in adult patients with brain tumours. Lancet Neurol 3:159–168. doi:10.1016/S1474-4422(04)00680-5
16. Tu JV, Willison DJ, Silver FL, Fang J, Richards JA, Laupacis A, Kapral MK, for the investigators in the registry of the Canadian Stroke Network (2004) Impracticability of informed consent in the Registry of the Canadian Stroke Network. N Engl J Med 350:1414–1421. doi:10.1056/NEJMsa031697
17. Westphal M, Hilt DC, Bortey E, Delavault P, Olivares R, Warnke PC, Whittle IR, Jääskeläinen J, Ram Z (2003) A phase 3 trial of local chemotherapy with biodegradable carmustine (BCNU) wafers (Gliadel wafers) in patients with primary malignant glioma. Neuro-oncol 5(2):79–88. doi:10.1215/15228517-5-2-79
18. Whittle IR, Thomson AM, Taylor R (1998) The effects of resective surgery for left sided intracranial tumours on language function: a prospective study. Lancet 351:1014–1018. doi:10.1016/S0140-6736(97)08295-0
19. Whittle IR, Lyles S, Walker M (2003) Gliadel therapy given for first resection of malignant glioma: a single centre study of the potential use of Gliadel. Br J Neurosurg 17(4):352–354. doi:10.1080/02688690310001601252
20. Whittle IR (2004) The dilemma of low grade glioma. J Neurol Neurosurg Psychiatry 75(Suppl II):ii31–ii36. doi:10.1136/jnnp.2004.040501
21. Whittle IR, Short DM, Deighton RF, McCulloch J (2007) Proteomic analysis of gliomas. Br J Neurosurg 21(6):576–582. doi:10.1080/02688690701721691
22. Wiemann CM, Chacko MR, Tucker JC, Velasquez MM, Smith PB, DiClemente RJ, von Sternberg K (2005) Enhancing recruitment and retention of minority young women in community-based clinical research. J Pediatr Adolesc Gynecol 18:403–407. doi:10.1016/j.jpag.2005.09.006
23. Williams B, Irvine L, McGinnis AR, McMurdo ME, Crombie IK (2007) When "no" might not quite mean "no"; the importance of informed and meaningful non-consent: results from a survey of individuals refusing participation in a health-related research project. BMC Health Serv Res 7:59–69. doi:10.1186/1472-6963-7-59
24. Wilson-Barnett J, Griffiths P (2002) The disappearing sample: researcher and research ability. Int J Nurs Stud 39:365–367. doi:10.1016/S0020-7489(01)00043-8
25. Zbinden L, Pandey R, Schwindack C, Deary I, Whittle IR (2006) Information processing in patients with intracranial tumours: a preliminary evaluation of the clinical utility of inspection time. Br J Neurosurg 20(4):203–208. doi:10.1080/02688690600852258

Comments

This is a clearly written, timely, important, wise, and honest European study on the recruitment bias in clinical neurosurgical research. This and similar ones should be published in high impact journals but would be rejected as nasty reminders that much cheered multicenter studies (ISAT, ISUIA, et al) and large multiethnic patient cohorts gathered from here and there - immortalized in high impact journals and eternally referred and discussed thereafter which nicely increases impact factors - represent a haphazard portion of our daily spectrum of neurosurgical diseases in our unselected and ageing European catchment populations.

Juha E Jääskeläinen
Kuopio, Finland

Returners vs Non-Returners: Comparison of Baseline Scores

U.1 Overview of Analysis Procedure

Before examining the effects of surgical intervention on cognitive test performance, the baseline test scores of those participants in each group (brain tumour and spinal surgery control) who completed session 2 follow-up were compared with those who did not complete follow-up at session 2 using univariate general linear modelling (analysis of covariance). This was done in order to determine whether those participants who were not followed up post-operatively were those patients who had the most severe impairment prior to surgery. Participants were classified as 'returners' if they completed session 2 testing and as 'non-returners' if they were not tested at session 2. The data was analysed separately for each of the participant groups. Follow-up status (i.e. returner or non-returner) and sex were entered as fixed effects (between-subjects factors) in the models. Age and NART score were included as covariates for the reasons described in Chapter 5.2.1.

In each of the sections that follow, mean scores for the returners and non-returners in the brain tumour group and the spinal control group are presented with standard deviations (SD) in brackets. The results of the general linear modelling analysis are then reported. Estimated marginal means (adjusted for age and NART score) are reported for the returner and non-returner groups and for the male and female participants in each group where applicable. Since the number of non-returners in the healthy group was too small to allow for any meaningful comparisons (n = 4), only comparisons of the returners and non-returners in each of the two surgical groups (brain tumour and spinal surgery control) are presented below. An overview of the comparisons of the returners and non-returners in each of the groups is shown in Table U.1.

U.2 Demographic comparisons

Sex

Brain Tumour Group

Thirty-two members of the returner group were male and 32 were female. Twenty-eight members of the non-returner group were male and 26 were female. Pearson chi-square analysis showed there were no significant differences between the two groups with respect to sex, $\chi^2(1) = 0.040$, $p = 0.841$.

Spinal Surgery Group

In the spinal surgery group, 34 returners were male and 33 were female. In the non-returner group, there were 5 males and 13 females. Pearson chi square analysis showed there were no significant differences between the two groups with respect to sex, $\chi^2(1) = 3.014$, $p = 0.083$.

Age

Brain Tumour Group

The mean age of the returners in the brain tumour group was 49.6 years (SD 14.2) and the mean age of the non-returners was 50.5 years (SD 11.5). The independent samples t-test revealed no significant differences in the age of the returners and non-returners, $t(116) = -0.398$, $p = 0.691$.

Spinal Surgery Group

The mean age of the returners in the spinal control group was 46.9 (SD 10.7) and the mean age of the non-returners was 51.1 (SD 12.9). The independent samples t-test revealed no significant differences in the age of the returner and non-returners, $t(83) = -1.433$, $p = 0.156$.

Handedness

Brain Tumour Group

In the returner group, 56 patients were right handed and 8 were left handed. In the non-returner group, 46 patients were right handed and 7 were left handed. Pearson chi square analysis revealed no significant difference between the two groups with respect to handedness, $\chi^2(1) = 0.013$, $p = 0.909$.

Spinal Surgery Group

In the returner group, 60 patients were right handed, 6 were left handed and 1 patient was ambidextrous. In the non-returner group there were 18 right handed patients and no left handed patients. Pearson chi square analysis revealed no significant difference between the returners and non-returners with respect to handedness, $\chi^2(1) = 2.049$, $p = 0.359$.

Highest Levels of Qualification Achieved

Brain Tumour Group

Pearson chi square analysis of the highest levels of qualification achieved by participants in each of the returner and non-returner groups revealed no significant difference between the two groups, $\chi^2(6) = 8.307$, $p = 0.216$.

Spinal Surgery Group

Pearson chi square analysis of the highest levels of qualification achieved by participants in the returner and non-returner groups revealed no significant difference between then two groups, $\chi^2(5) = 1.439$, $p = 0.920$.

National Adult Reading Test Score

Brain Tumour Group

The mean NART score for the returners in the brain tumour group was 18.8 (SD 8.5) and for the non-returners was 22.8 (SD 10.0). The independent samples t-test revealed a significant difference between the NART scores of the two groups, $t(114) = -2.338$, $p = 0.021$. Although there were no significant differences between the highest levels of qualification achieved in the returners and non-returners, the non-returners in the brain tumour group had significantly higher (poorer) NART scores than the returners.

Spinal Control Group

The returners in the spinal control group had a mean NART score of 19.7 (SD 9.2) and the non-returners had a mean score of 21.7 (SD 10.0). The independent samples t-test showed there to be no significant difference between the NART scores of the two groups, $t(81) = -0.880$, $p = 0.382$.

Tumour Type

Brain Tumour Group

In the returner group, 26 patients had a high-grade glioma; 16 had a low-grade glioma; 9 had a metastasis; 10 had a meningioma and 3 had 'other' tumour types. The non-returner group comprised 25 high-grade glioma patients; 7 low-grade glioma patients; 5 patients with a metastasis; 7 with a meningioma and 8 patients with 'other' tumour types. Pearson chi-square analysis was carried out to compare the histological diagnoses of those participants who were returners and those who were non-returners at session 2. There was no significant difference between the two groups, $\chi^2(5) = 8.701$, $p = 0.122$.

U.3 Baseline Test Comparisons

Inspection Time: All Inspection Time Scores

Brain Tumour Group

Sixty two patients in the brain tumour cohort completed inspection time testing at baseline and session 2 ('returners') and 50 patients in the brain tumour cohort completed baseline testing but did not return for session 2 follow-up ('non-returners'). The mean baseline inspection time score for the returner group was 116.6 (SD 21.2) and for the non-returner group was 111.5 (SD 17.9).

General linear modelling, with inspection time total (i.e. score /150) as the dependent variable revealed a significant main effect of the covariates age, $F(1,106) = 37.711$, $p < 0.001$, partial $\eta^2 = 0.262$; and NART score, $F(1,106) = 6.563$, $p = 0.012$, partial $\eta^2 = 0.058$, in the model. Older patients, and those with higher (i.e. poorer) NART scores tended to perform less well on inspection time at baseline. There was no significant main effect of sex in the model, $F(1,106) = 2.621$, $p = 0.108$, partial $\eta^2 = 0.024$. There was no significant main effect of follow-up status (returner or non-returner) on baseline inspection time scores, in the model that included the effects of age and NART score, $F(1,106) = 0.494$, $p = 0.484$, partial $\eta^2 = 0.005$. The interaction between sex and follow-up status was not significant, $F(1,106) = 0.639$, $p = 0.426$, partial $\eta^2 = 0.006$.

The estimated marginal mean baseline inspection time scores for the brain tumour group, adjusted for age and NART score, were 115.4 (SE 2.2) for the returners and 113.1 (SE 2.4) for the non-returners. Therefore, there was no significant difference between the returners and non-returners in the brain tumour group on baseline inspection time performance when all inspection time data was included in the model.

Spinal Surgery Group

Sixty-six patients in the spinal cohort completed inspection time testing at baseline and session 2 (returners) and a further 16 were classified as 'non-returners' since they were not tested on inspection time at session 2.

The mean baseline inspection time score for the returner group in the spinal control cohort was 122.2 (SD 13.8), and the mean score for the non-returner group was 118.0 (SD 13.0).

General linear modelling was carried out with inspection time total entered as the dependent variable. There was a significant main effect of the covariate age in the model, $F(1,76) = 11.836$, $p = 0.001$, partial $\eta^2 = 0.135$. The covariate NART score also had a significant main effect in the model, $F(1,76) = 9.041$, $p = 0.004$, partial $\eta^2 = 0.106$. Older patients and those who scored less well on the NART measure tended to have poorer inspection time scores. There was no significant main effect of sex, $F(1,76) = 0.768$, $p = 0.384$, partial $\eta^2 = 0.010$. Follow-up status (returner or non-returner) had no significant main effect on baseline inspection time scores in the model that included the effects of age and NART score, $F(1,76) = 0.001$, $p = 0.974$, partial $\eta^2 < 0.001$. There was no significant interaction between sex and follow-up status, $F(1,76) = 2.693$, $p = 0.105$, partial $\eta^2 = 0.034$.

The estimated marginal mean baseline inspection time scores for the spinal group, adjusted for age and NART score, were 121.7 (SE 1.5) for the returner group and 121.9 (SE 3.4) for the non-returner group. This shows that, when all baseline inspection time scores were included in the model, there was no significant difference in inspection time performance between those participants in the spinal control group who completed session 2 follow-up and those who completed baseline testing only in the spinal group.

Inspection Time: Valid Inspection Time Scores

Brain Tumour Group

Fifty-three patients in the brain tumour cohort had 'valid' inspection time scores at baseline and follow-up scores at session 2 ('returners') and 43 patients in the brain tumour group had valid inspection time scores at baseline but did not complete session 2 testing ('non-returners'). Valid inspection time scores are scores from those patients who scored at least 17/20 on the two longest durations, as described in chapter 2.2.

The mean baseline score for the brain tumour returner group with valid baseline inspection time data was 123.3 (SD 13.1) and for the non-returner group was 116.2 (SD 13.5).

General linear modelling revealed a significant main effect of the covariate age on valid baseline inspection time scores, $F(1,90) = 23.888$, $p < 0.001$, partial $\eta^2 = 0.210$. Older participants tended to have lower inspection time scores. The covariate NART score had no significant main effect on the model when only valid inspection time scores were included, $F(1,90) = 2.123$, $p = 0.149$, partial $\eta^2 = 0.023$. The effect of sex approached the conventional level of statistical significance in the model, $F(1,90) = 3.918$, $p = 0.051$, partial $\eta^2 = 0.042$. Follow-up status (returner or non-returner) had a significant main effect in the model that included the effects of age and NART score on valid baseline inspection time scores, $F(1,90) = 4.855$, $p = 0.030$, partial $\eta^2 = 0.051$. There was a significant interaction between sex and follow-up status, $F(1,90) = 4.288$, $p = 0.041$, partial $\eta^2 = 0.045$.

The estimated marginal mean baseline inspection time scores for brain tumour patients with valid data, adjusted for age and NART score, were 122.6 (SE 1.6) for the returner group and 117.3 (SE 1.8) for the non-returner group. Therefore, in this group of patients with valid baseline scores, those brain tumour patients who were returners performed significantly better at baseline than those who did not complete follow-up (non-returners).

The estimated marginal mean score for the female returners was 122.5 (SE 2.2) and for the female non-returners was 112.3 (SE 2.4). The corresponding mean score for the male returners was 122.8 (SE 2.4) and for the male non-returners was 122.2 (SE 2.7). The female non-returners performed significantly worse than the female returners on inspection time at baseline, when only valid data was included in the analysis.

Spinal Surgery Group

When only 'valid' baseline inspection time scores were included, 62 members of the spinal control group were classified as 'returners' and 16 patients had valid baseline inspection time scores but did not complete session 2 testing ('non-returners').

The mean baseline score for the returner group was 124.7 (SD 10.1) and for the non-returner group was 118.0 (SD 13.0).

General linear modelling revealed a significant main effect of age on baseline inspection time scores when only valid data was entered into the model, $F(1,72) = 17.767$, $p < 0.001$, partial $\eta^2 = 0.198$. Older patients were again more likely to perform less well on this test. The covariate NART score had no significant main effect on baseline inspection time scores on this occasion, $F(1,72) = 2.558$, $p = 0.114$, partial $\eta^2 = 0.034$. Sex also had no significant main effect in the model, $F(1,72) = 1.366$, $p = 0.246$, partial $\eta^2 = 0.019$. There was no significant main effect of follow-up status (returner or non-returner) on baseline valid inspection time scores in the model that included the effects of age and NART score, $F(1,72) = 0.822$, $p = 0.368$, partial $\eta^2 = 0.011$. There was however a significant interaction between follow-up status and sex in the model, $F(1,72) = 5.322$, $p = 0.024$, partial $\eta^2 = 0.069$.

The estimated marginal mean valid baseline inspection time scores, adjusted for age and NART score were 124.2 (SE 1.2) for the returner group and 121.6 (SE 2.6) for the non-returner group. There was no significant difference between the performance of the returners and non-returners with valid baseline inspection time scores in the spinal group.

The estimated marginal mean score for the female returners was 125.8 (SE 1.7), compared with 116.6 (SE 2.9) for the female non-returners. The corresponding mean score for the male returners was 122.6 (SE 1.7) and for the male non-returners was 126.5 (SE 4.3). Therefore, the significant interaction between follow-up status and sex suggests that female non-returners scored significantly worse in inspection time at baseline than the female returners when only valid scores were included.

Rey Auditory Verbal Learning Test

Brain Tumour Group

There were 51 brain tumour patients who completed the Rey Auditory Verbal Learning Test (RAVLT) at baseline and also completed session 2 ('returners') and 36 brain tumour patients who completed this test at baseline but did not complete the follow-up testing session ('non-returners').

The returners scored a mean of 62.7 (SD 19.0) on this measure at baseline and the non-returners score a mean of 60.3 (SD 15.4).

General linear modelling, with RAVLT total score entered as the dependent variable, revealed a significant main effect of the covariate age on test score, $F(1,81) = 24.731$, $p < 0.001$, partial $\eta^2 = 0.234$. NART score also had a significant main effect in the model, $F(1,81) = 15.382$, $p < 0.001$, partial $\eta^2 = 0.160$. Participants with better (lower) NART scores and younger participants were significantly more likely to have higher (better) RAVLT scores. Sex also had a significant main effect in the model, $F(1,81) = 4.897$, $p = 0.030$, partial $\eta^2 = 0.057$. Follow-up status (returner or non-returner) had no significant main effect on RAVLT scores in the model that included the effects of age and NART score, $F(1,81) = 0.503$, $p = 0.480$, partial $\eta^2 = 0.006$. There was also no significant interaction between sex and follow-up status, $F(1,81) = 0.469$, $p = 0.496$, partial $\eta^2 = 0.006$.

Estimated marginal mean scores, adjusted for age and NART score were 60.9 (SE 2.1) for the returner group and 63.3 (SE 2.5) for the non-returner group. There was

no significant difference between the performance of the returners and non-returners in the brain tumour group on the RAVLT at baseline.

Spinal Surgery Group

There were 60 patients in the spinal group who completed the RAVLT at baseline and also completed session 2 testing ('returners'). Fourteen members of the spinal group completed this test at baseline but did not complete any follow-up testing ('non-returners').

The mean RAVLT score for the returners in the spinal group was 72.5 (SD 15.6) and for the non-returners was 64.8 (SD 15.5).

General linear modelling revealed a significant main effect of the covariates age, $F(1,68) = 36.006$, $p < 0.001$, partial $\eta^2 = 0.346$; and NART score, $F(1,68) = 19.868$, $p < 0.001$, partial $\eta^2 = 0.226$ in the model. Older patients, and those with higher (poorer) NART scores were significantly more likely to have lower scores on the RAVLT. There was no significant main effect of sex in the model, $F(1,68) = 0.642$, $p = 0.426$, partial $\eta^2 = 0.009$. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of age and NART score, $F(1,68) = 0.152$, $p = 0.698$, partial $\eta^2 = 0.002$. There was however a significant interaction between follow-up status and sex in the model, $F(1,68) = 4.858$, $p = 0.031$, partial $\eta^2 = 0.067$.

Estimated marginal mean RAVLT scores, adjusted for age and NART scores, for the spinal control group were 71.5 (SE 1.5) for the returners and 70.1 (SE 3.3) for the non-returners. There was therefore no significant overall difference between the performance of the returners and the non-returners on the RAVLT.

The estimated marginal mean scores for the female returners was 74.1 (SE 2.1) and for the female non-returners, was 64.6 (SE 4.0). The male returners had a corresponding mean score of 69.0 (SE 2.2) and the male non-returners had a mean of 64.6 (SE 4.0). It would therefore appear that the female non-returners performed significantly worse on the RAVLT than the female returners.

Trail Making Test Part B

Brain Tumour Group

Fifty-three patients in the brain tumour group completed the Trail Making Test Part B and also completed session 2 ('returners'). Forty-three brain tumour patients were classified as 'non-returners' who completed the test at baseline but did not participate in session 2 follow-up.

General linear modelling, with Trail Making Test Part B score entered as the dependent variable, revealed a significant main effect of the covariate age in the model, $F(1,90) = 19.000$, $p < 0.001$, partial $\eta^2 = 0.174$. Younger patients were significantly more likely to be faster to complete the test than older participants. The effect of the covariate NART score was also significant in the model, $F(1,90) = 17.894$, $p < 0.001$, partial $\eta^2 = 0.166$. Those participants in the brain tumour group who had better (lower) NART scores were also significantly more likely to have better (faster) scores on the Trail Making Test Part B. Sex had no significant main effect in the model, $F(1,90) = 0.075$, $p = 0.784$, partial $\eta^2 = 0.001$. The effect of follow-up status (returner or non-returner) approached the conventional level of statistical significance in the model that included the effects of age and NART score, $F(1,90) = 3.375$, $p = 0.069$, partial $\eta^2 = 0.036$. There was no significant interaction between follow-up status and sex in the model, $F(1,90) = 0.767$, $p = 0.383$, partial $\eta^2 = 0.008$.

The estimated marginal mean scores, adjusted for age and NART score, were 90.1 (SE 3.7) for the returner group and 100.5 (SE 4.2) for the non-returner group on the TMT Part B. Therefore, in the brain tumour group, the returners had faster scores than the non-returners on this test. However, the difference failed to reach the conventional level of statistical significance.

Spinal Surgery Group

Sixty-five patients in the spinal control group completed the Trail Making Test Part B at baseline and also took part in session 2 testing ('returners'). Seventeen spinal control patients were classified as 'non-returners'.

The returners in the spinal group had a mean score of 79.8 (SD 25.6) and the non-returners had a mean score of 86.6 (SD 16.2).

General linear modelling, with Trail Making Test Part B score entered as the dependent variable, revealed a significant main effect of the covariates age, $F(1,76) = 21.199$, $p < 0.001$, partial $\eta^2 = 0.218$; and NART score, $F(1,76) = 21.082$, $p < 0.001$, partial $\eta^2 = 0.217$, in the model. Again, those patients in the spinal group who were older or had higher (poorer) scores on the NART were significantly more likely to take longer to complete this test. There was no significant main effect of sex on test scores in this model, $F(1,76) = 0.070$, $p = 0.792$, partial $\eta^2 = 0.001$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,76) = 0.003$, $p = 0.955$, partial $\eta^2 < 0.001$. The interaction between sex and follow-up status was also non-significant, $F(1,76) = 0.039$, $p = 0.843$, partial $\eta^2 < 0.001$.

The estimated marginal mean scores on the TMT Part B, adjusted for age and NART score, were 81.3 (SE 2.5) for the spinal group returners and 80.9 (SE 5.5) for the spinal group non-returners. Therefore, there was no significant difference between the performance of the returners and non-returners in the spinal group on this test.

Verbal Fluency

Brain Tumour Group

Forty-nine patients in the brain tumour group were classified as returners who completed the verbal fluency test at baseline and also completed session 2 testing and 42 patients were 'non-returners' who completed the test at baseline only.

The mean verbal fluency total score for the returners in the brain tumour group was 31.4 (SD 13.6) and for the non-returners was 25.5 (SD 9.2).

General linear modelling, with total verbal fluency score as the dependent variable, revealed a significant main effect of the covariates age, $F(1,85) = 8.589$, $p = 0.004$, partial $\eta^2 = 0.092$; and NART score, $F(1,85) = 38.086$, $p < 0.001$, partial $\eta^2 = 0.309$, in the model. Participants who performed better on the NART and older participants tended to perform better on the test. Sex had no significant main effect in the model, $F(1,85) = 2.560$, $p = 0.113$, partial $\eta^2 = 0.029$. Follow-up status had no significant main effect in the model that included the effects of age and NART score, $F(1,85) = 2.275$, $p = 0.135$, partial $\eta^2 = 0.026$. The interaction between sex and follow-up status was significant in the model, $F(1,85) = 13.267$, $p < 0.001$, partial $\eta^2 = 0.135$.

The estimated marginal mean score, adjusted for age and NART, for the returner group on the verbal fluency measure was 30.2 (SD 1.4) and for the non-returner group was 27.1 (SD 1.5). There was no significant overall difference between the two follow-up groups on this measure at baseline.

The estimated marginal mean score for the female returners was 35.8 (SE 2.0) and for the female non-returners was 24.8 (SE 2.1). The corresponding mean for the male returners was 24.6 (SE 2.0) and for the male non-returners was 29.4 (SE 2.3). Therefore, in the brain tumour group, the female returners performed significantly better than the female non-returners on the verbal fluency measure.

Spinal Surgery Group

In the spinal control group, 64 patients were classified as returners who completed the verbal fluency test at baseline and also completed session 2 follow-up. Fourteen spinal patients were classified as non-returners who completed verbal fluency testing at baseline but did not complete session 2 follow-up.

The mean verbal fluency score for the returners in the spinal group was 35.3 (SD 13.6) and for the non-returners was 33.4 (SD 11.0).

General linear modelling with verbal fluency total score for each spinal group patient, revealed that the effect of the covariate age approached the conventional level of statistical significance, $F(1,72) = 3.807$, $p = 0.055$, partial $\eta^2 = 0.050$. This suggests that older patients tended to perform better on the verbal fluency test than younger patients. The effect of the covariate NART was significant in the model, $F(1,72) = 33.966$, $p < 0.001$, partial $\eta^2 = 0.321$. Those patients with better scores on the NART tended to perform better on the verbal fluency measure. The effect of sex was significant in the model, $F(1,72) = 6.018$, $p = 0.017$, partial $\eta^2 = 0.077$. Female patients performed significantly better than male patients. The effect of follow-up status was not significant in the model that included the effects of age and NART score, $F(1,72) = 0.454$, $p = 0.503$, partial $\eta^2 = 0.006$. There was also no significant interaction between age and follow-up status, $F(1,72) = 0.623$, $p = 0.433$, partial $\eta^2 = 0.009$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the spinal control group was 35.2 (SE 1.4) and for the non-returners was 32.9 (SE 3.1). There was no significant difference between the returners and non-returners in the spinal group on the verbal fluency test.

Digit Symbol Coding

Brain Tumour Group

Sixty-two members of the brain tumour group completed this measure at baseline and also completed session 2 follow-up (returners) and 49 patients had baseline digit symbol coding data but did not complete follow-up (non-returners).

The returner group had a mean score of 61.5 (SD 22.4) on digit symbol coding at baseline and the non-returner group had a mean score of 49.9 (SD 20.3).

General linear modelling, with baseline digit symbol coding score entered as the dependent variable, revealed a significant main effect of the covariate age on the test scores, $F(1,105) = 42.265$, $p < 0.001$, partial $\eta^2 = 0.287$. Older patients tended to perform less well on the digit symbol coding measure. The covariate NART score

also had a significant main effect in the model, $F(1,105) = 23.547$, $p < 0.001$, partial $\eta^2 = 0.183$. Patients with higher (poorer) NART scores were more likely to have lower (poorer) digit symbol coding scores. There was no significant main effect of sex in the model, $F(1,105) = 0.485$, $p = 0.488$, partial $\eta^2 = 0.005$. There was a significant main effect of follow-up status in the model that included the effects of the covariates, $F(1,105) = 5.164$, $p = 0.025$, partial $\eta^2 = 0.047$. The interaction between follow-up status and sex was not significant, $F(1,105) = 1.307$, $p = 0.256$, partial $\eta^2 = 0.012$.

The estimated marginal mean score for the returners in the brain tumour group was 59.6 (SE 2.2) and for the non-returners was 52.0 (SE 2.5). The non-returners in the brain tumour group performed significantly less well on the digit symbol coding measure than those who returned to take part in session 2.

Spinal Surgery Group

There were 66 spinal patients who were classified as 'returners' who completed digit symbol coding at baseline and also returned for session 2 follow-up. There were 17 spinal patients who completed the test at baseline but did not complete session 2 ('non-returners').

The mean baseline digit symbol coding score for the returner group was 69.4 (SD 17.2) and for the non-returner group was 59.6 (SD 13.9).

General linear modelling, with digit symbol coding score at baseline entered as the dependent variable, revealed a significant main effect of the covariate age in the model, $F(1,77) = 38.621$, $p < 0.001$, partial $\eta^2 = 0.334$. Older patients were significantly more likely to have lower scores on the measure. The effect of the covariate NART was also significant, $F(1,77) = 27.503$, $p < 0.001$, partial $\eta^2 = 0.263$. Patients with better NART scores were significantly more likely to perform better on the digit symbol coding test. The effect of sex was not significant in the model, $F(1,77) = 1.212$, $p = 0.274$, partial $\eta^2 = 0.016$. There was also no significant main effect of follow-up status (returner or non-returner) on baseline digit symbol coding

scores in the model that included the effects of age and NART score, $F(1,77) = 1.066$, $p = 0.305$, partial $\eta^2 = 0.014$. There was no significant interaction between sex and follow-up status in the model, $F(1,77) = 1.584$, $p = 0.212$, partial $\eta^2 = 0.020$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the brain tumour group on the digit symbol coding test was 68.2 (SD 1.6) and for the non-returners was 64.3 (SD 3.4). The difference between the baseline scores was not significant.

Letter-Number Sequencing

Brain Tumour Group

There were 47 'returners' in the brain tumour group who completed letter-number sequencing at baseline and who also returned for follow-up testing at session 2. Forty-two brain tumour patients were 'non-returners' who completed the test at baseline but did not complete session 2 follow-up.

The mean letter-number sequencing score for the returners in the brain tumour group was 9.6 (SD 3.4) and for the non-returners was 9.2 (SD 3.3).

General linear modelling, with baseline letter-number sequencing test score as the dependent variable, revealed a significant main effect of age on test scores, $F(1,83) = 19.548$, $p < 0.001$, partial $\eta^2 = 0.191$. Older patients were significantly more likely to have lower scores on this test. The covariate NART score also had a significant main effect in the model, $F(1,83) = 19.364$, $p < 0.001$, partial $\eta^2 = 0.189$. Those participants with higher (worse) scores on the NART were more likely to have lower (worse) scores on letter-number sequencing. There was no significant main effect of sex in the model, $F(1,83) = 0.235$, $p = 0.629$, partial $\eta^2 = 0.003$. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of age and NART score, $F(1,83) = 0.726$, $p = 0.397$, partial $\eta^2 = 0.009$. There was no significant interaction between sex and follow-up status, $F(1,83) = 1.545$, $p = 0.217$, partial $\eta^2 = 0.018$.

The estimated marginal mean score, adjusted for age and NART score, for this measure was 9.2 (SE 0.4) for the returners in the brain tumour group and was 9.7 (SE 0.4) for the non-returners. There was no significant difference between the performance of the returners and the non-returners on the letter-number sequencing measure.

Spinal Surgery Group

In the spinal group, there were 63 patients who completed the letter-number sequencing test at baseline and returned for session 2 follow-up (returners). Fifteen spinal patients completed the test at baseline but did not complete session 2 follow-up (non-returners).

The mean score for the returners in the spinal group on this test was 11.3 (SD 2.7) and for the non-returners was 10.9 (SE 2.5).

General linear modelling, with letter-number sequencing test score as the dependent variable, revealed a significant main effect of the covariates age, $F(1,72) = 23.436$, $p < 0.001$, partial $\eta^2 = 0.246$; and NART score, $F(1,72) = 16.779$, $p < 0.001$, partial $\eta^2 = 0.189$ in the model. Older patients and those with poorer scores on the NART were significantly more likely to have poorer letter-number sequencing test scores. There was no significant main effect of sex in the model, $F(1,72) = 0.283$, $p = 0.596$, partial $\eta^2 = 0.004$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,72) = 0.265$, $p = 0.596$, partial $\eta^2 = 0.004$. There was no significant interaction between sex and follow-up status, $F(1,72) = 0.002$, $p = 0.961$, partial $\eta^2 < 0.001$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the spinal group on this test was 11.1 (SE 0.3) and for the non-returners was 11.5 (SE 0.6). The returners did not perform significantly differently from the non-returners on the letter-number sequencing test at baseline.

Williams Delayed Recall Test (EFIT)

Brain Tumour Group

Sixty-three patients in the brain tumour group completed the Williams Delayed Recall Test (WDRT) at baseline and also completed session 2 follow-up ('returners'), compared with 52 patients who completed the test at baseline only, and did not complete session 2 ('non-returners').

The mean score for the returners on this measure was 8.8 (SD 5.2) and for the non-returners was 7.5 (SD 4.9).

General linear modelling, with WDRT score as the dependent variable, revealed a significant main effect of age on test scores, $F(1,109) = 14.550$, $p < 0.001$, partial $\eta^2 = 0.118$. Older patients were significantly more likely to have higher (poorer) scores on this test. The covariate NART score also had a significant main effect in the model, $F(1,109) = 4.102$, $p = 0.045$, partial $\eta^2 = 0.036$. Patients who had lower (better) NART scores were significantly more likely to have lower (better) scores on this test. There was no significant main effect of sex in the model, $F(1,109) = 1.521$, $p = 0.220$, partial $\eta^2 = 0.014$. The effect of follow-up status approached a conventional level of significance in the model that included the effects of age and NART score, $F(1,109) = 3.889$, $p = 0.051$, partial $\eta^2 = 0.034$. There was no significant interaction between the sex of the participant and follow-up status, $F(1,109) = 1.379$, $p = 0.243$, partial $\eta^2 = 0.012$.

The estimated marginal mean score, adjusted for age and NART score, on the WDRT, was 9.0 (SE 0.6) for the returners in the brain tumour group and for the non-returner group was 7.2 (SE 0.7). The returner group performed better than the non-returner group on this test at baseline, and the difference between the groups approached the conventional level of statistical significance.

Spinal Surgery Group

A total of 66 patients in the spinal control group were classified as 'returners' who completed the WDRT at baseline and also completed session 2 testing. Seventeen

patients were classed as ‘non-returners’ who completed the test at baseline but did not complete follow-up.

The mean score for the returners in the spinal group on the WDRT was 5.0 (SD 3.7) and for the non-returners in the spinal group was 6.6 (SD 3.8).

WDRT score at baseline was the dependent variable in the general linear modelling analysis. There was a significant main effect of age in the model, $F(1,77) = 5.715$, $p = 0.019$, partial $\eta^2 = 0.069$. Older participants were significantly more likely to have higher (poorer) scores on the test. There was also a significant main effect of NART score on WDRT scores, $F(1,77) = 4.337$, $p = 0.041$, partial $\eta^2 = 0.053$. Lower (better) NART scores were significantly related to lower (better) WDRT scores. There was no significant main effect of sex in the model, $F(1,77) = 0.229$, $p = 0.634$, partial $\eta^2 = 0.003$. Follow-up status (returner or non-returner) had no significant effect in the model that included the effects of age and NART score, $F(1,77) = 0.646$, $p = 0.424$, partial $\eta^2 = 0.008$. There was no significant interaction between participant sex and follow-up status, $F(1,77) = 0.777$, $p = 0.381$, partial $\eta^2 = 0.010$.

The estimated marginal mean score, adjusted for age and NART score, on the WDRT was 5.1 (SD 0.4) for the returner group and 6.0 (SD 1.0) for the non-returner group. There was no significant difference between the performance of the returners and the non-returners in the spinal group on this test at baseline.

Nine Hole Peg Test – Right Hand

Brain Tumour Group

A total of 63 patients in the brain tumour group completed the right hand Nine Hole Peg Test (NHPT) at baseline and also completed session 2 follow-up (‘returners’) and 52 patients completed the test at baseline only and did not complete session 2 (‘non-returners’).

The mean score on the right hand NHPT for the returners in the brain tumour group was 13.8 (SD 2.9) and for the non-returners was 15.8 (SD 5.8).

Right hand NHPT score was the dependent variable in the general linear modelling analysis. There was a significant main effect of the covariate age in the model, $F(1,109) = 13.095$, $p < 0.001$, partial $\eta^2 = 0.107$. Older participants were more likely to perform less well on this measure. The effect of the covariate NART score was also significant in the model, $F(1,109) = 9.743$, $p = 0.002$, partial $\eta^2 = 0.082$. Those participants who scored less well on the NART were more likely to have poorer scores on the right hand NHPT. The effect of sex was not significant in the model, $F(1,109) = 0.754$, $p = 0.387$, partial $\eta^2 = 0.007$. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of age and NART score, $F(1,109) = 3.003$, $p = 0.086$, partial $\eta^2 = 0.027$. There was no significant interaction between sex and follow-up status, $F(1,109) = 0.002$, $p = 0.964$, partial $\eta^2 < 0.001$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the brain tumour group on the right hand NHPT was 14.1 (SE 0.5) and for the non-returners was 15.5 (SE 0.6). The non-returners were slower to complete this test than the returners, however, the difference between the two groups was not significant.

Spinal Surgery Group

Sixty-six patients in the spinal group were classed as ‘returners’ who completed the right hand NHPT at baseline and also completed session 2 follow-up. There were 17 ‘non-returners’ in the spinal group who completed the test at baseline but did not complete follow-up.

The mean score for the returners was 13.0 (SD 2.9) and for the non-returners was 12.9 (SD 1.8).

General linear modelling, with right hand NHPT score as the dependent variable, revealed a significant main effect of age on test scores, $F(1,77) = 9.875$, $p = 0.002$, partial $\eta^2 = 0.114$. Older patients were significantly more likely to take longer to complete the test at baseline. The effect of the covariate NART score was not

significant in the model, $F(1,77) = 3.208$, $p = 0.077$, partial $\eta^2 = 0.040$. There was no significant main effect of sex in the model, $F(1,77) = 0.133$, $p = 0.716$, partial $\eta^2 = 0.002$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,77) = 0.628$, $p = 0.430$, partial $\eta^2 = 0.008$. The interaction between sex and follow-up status was not significant, $F(1,77) = 0.285$, $p = 0.595$, partial $\eta^2 = 0.004$.

The estimated marginal mean score, adjusted for age and NART score, on the right hand nine hole peg test was 13.1 (SE 0.3) for the returners in the spinal control group and was 12.5 (SE 0.7) for the non-returners. There was no significant difference between the returners and the non-returners in terms of performance on this test at baseline.

Nine Hole Peg Test

Brain Tumour Group

Sixty-three patients in the brain tumour group completed the left hand Nine Hole Peg Test (NHPT) at baseline and also completed session 2 follow-up ('returners') and 49 patients completed the test at baseline but did not complete session 2 ('non-returners').

The mean score for the returners on the left hand NHPT was 15.5 (SD 3.9) and for the non-returners was 15.3 (SD 3.0).

General linear modelling, with left hand NHPT score as the dependent variable revealed a significant main effect of the covariate age in the model, $F(1,106) = 21.188$, $p < 0.001$, partial $\eta^2 = 0.167$. Older participants were significantly more likely to take longer to complete the left hand NHPT. The effect of the covariate NART was not significant in the model, $F(1,106) = 2.119$, $p = 0.148$, partial $\eta^2 = 0.020$. Sex had no significant main effect in the model, $F(1,106) = 2.864$, $p = 0.093$, partial $\eta^2 = 0.026$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,106) =$

0.362, p = 0.549, partial η^2 = 0.003. There was also no significant interaction between sex and follow-up status, F(1,106) = 0.008, p = 0.929, partial η^2 < 0.001.

The estimated marginal mean score, adjusted for age and NART, for the returners in the brain tumour group was 15.6 (SE 0.4) and for the non-returners was 15.2 (SE 0.5). There was no significant difference between the performance of the returners and the non-returners in the brain tumour group on the left hand NHPT at baseline.

Spinal Surgery Group

Sixty-six members of the spinal control group were 'returners' who completed the left hand nine hole peg test at baseline and also completed session 2 testing. Seventeen members of the spinal control group were 'non-returners' who completed the test at baseline but did not complete session 2 follow-up.

The mean score for the returners on the brain tumour group was 14.2 (SD 3.3) and for the non-returners was 14.8 (SD 4.8).

General linear modelling, with left hand NHPT score as the dependent variable, revealed a significant main effect of the covariate age on test scores, F(1,77) = 15.259, p < 0.001, partial η^2 = 0.165. Older participants tended to have slower scores on the test. The covariate NART score also had a significant main effect in the model, F(1,77) = 8.524, p = 0.005, partial η^2 = 0.100. Participants with higher (poorer) NART scores were more likely to take longer to complete the left hand NHPT. The effect of sex was not significant in the model, F(1,77) = 0.081, p = 0.776, partial η^2 = 0.001. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of age and NART score, F(1,77) = 0.222, p = 0.638, partial η^2 = 0.003. There was no significant interaction between sex and follow-up status, F(1,77) = 0.057, p = 0.812, partial η^2 = 0.001.

The estimated marginal mean score, adjusted for age and NART, for the returners in the brain tumour group was 14.4 (SE 0.4) and for the non-returners was 13.9 (SE 0.9). There was no significant difference between the performance of the returners and non-returners in the brain tumour group on the left hand nine hole peg test.

Timed Ten Metre Walk (EFIT)

Brain Tumour Group

Fifty-six members of the brain tumour group were 'returners' who completed the time ten metre walk at baseline and also completed session 2. There were 46 'non-returners' who completed the test at baseline but did not take part in session 2.

The returners scored a mean of 6.7 (SD 1.4) and the non-returners scored a mean of 7.3 (SE 2.1) on the timed ten metre walk at baseline.

General linear modelling, with timed ten metre walk as the dependent variable, revealed a significant main effect of age on test scores, $F(1,96) = 11.050$, $p = 0.001$, partial $\eta^2 = 0.103$. Older participants were significantly more likely to take longer to complete the time ten metre walk. The covariate NART score had no significant main effect in the model, $F(1,96) = 2.000$, $p = 0.160$, partial $\eta^2 = 0.020$. There was no significant effect of sex in the model, $F(1,96) = 1.335$, $p = 0.251$, partial $\eta^2 = 0.014$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,96) = 0.480$, $p = 0.490$, partial $\eta^2 = 0.005$. The interaction between sex and follow-up status was not significant, $F(1,96) = 0.188$, $p = 0.666$, partial $\eta^2 = 0.002$.

The estimated marginal mean, adjusted for age and NART score, for the returners in the brain tumour group on the timed ten metre walk, was 6.9 (SE 0.2) and for the non-returners was 7.1 (SE 0.3). There was no significant difference between the performance of the returners and non-returners on the time ten metre walk at baseline.

Spinal Surgery Group

There were 63 spinal control patients who completed the timed ten metre walk at baseline and also completed session 2 follow-up ('returners'). Fourteen spinal control patients were 'non-returners' who completed this test at baseline but did not take part in session 2.

The mean score for the returners in the spinal group on the timed ten metre walk was 7.9 (SD 2.3) and for the non-returners the mean score was 8.4 (SD 2.4).

General linear modelling, with timed ten metre walk score as the dependent variable, revealed a significant main effect of the covariate age, $F(1,71) = 8.344$, $p = 0.005$, partial $\eta^2 = 0.105$. Older patients were significantly more likely to take longer to complete the test than younger patients. There was no significant main effect of the covariate NART score, $F(1,71) = 1.247$, $p = 0.268$, partial $\eta^2 = 0.017$. The effect of sex was not significant in the model, $F(1,71) = 0.528$, $p = 0.470$, partial $\eta^2 = 0.007$. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of age and NART score, $F(1,71) = 0.030$, $p = 0.863$, partial $\eta^2 < 0.001$. The interaction between sex and follow-up status was not significant, $F(1,71) = 0.003$, $p = 0.955$, partial $\eta^2 < 0.001$.

The estimated marginal mean score on the timed ten metre walk, adjusted for age and NART, was 7.9 (SE 0.3) for the returners in the spinal control group and was 8.0 (SE 0.7) for the non-returners. There was no significant difference between the performance of the returners and the non-returners on the timed ten metre walk at baseline.

Hospital Anxiety and Depression Scale – Anxiety Score

Brain Tumour Group

There were 63 patients in the brain tumour group who completed the Hospital Anxiety and Depression Scale (HADS) at baseline and also completed session 2 follow-up ('returners') and 52 patients who completed the questionnaire at baseline but did complete session 2 testing ('non-returners').

The mean anxiety score for the returners was 7.9 (SD 4.6) and for the non-returners was 8.1 (SD 4.2).

General linear modelling, with anxiety scores as the dependent variable, revealed no significant main effect of either the covariate age, $F(1,109) = 0.088$, $p = 0.768$,

partial $\eta^2 = 0.001$; or NART score, $F(1,109) = 0.079$, $p = 0.779$, partial $\eta^2 = 0.001$. Sex had no significant main effect in the model, $F(1,109) = 0.816$, $p = 0.368$, partial $\eta^2 = 0.007$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of the covariates age and NART score, $F(1,109) = 0.043$, $p = 0.837$, partial $\eta^2 = 0.007$. The interaction between sex and follow-up status was not significant, $F(1,109) = 1.518$, $P = 0.221$, partial $\eta^2 = 0.014$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the brain tumour group was 7.9 (SE 0.6) and for the non-returners was 8.1 (SE 0.6). There was no significant difference in anxiety levels, as measured by the Hospital Anxiety and Depression Scale, between the returners and non-returners in the brain tumour group.

Spinal Surgery Group

There were 65 patients in the spinal control group who completed the Hospital Anxiety and Depression Scale at baseline and also completed session 2 follow-up ('returners') and 17 patients who completed the questionnaire at baseline but did not complete follow-up testing ('non-returners').

The mean anxiety score for the returners in the spinal control group was 8.1 (SD 4.0) and for the non-returners was 7.2 (SD 3.1).

General linear modelling, with anxiety scores at the dependent variable, revealed no significant main effect of either of the covariates age, $F(1,76) = 0.031$, $p = 0.860$, partial $\eta^2 < 0.001$; or NART score, $F(1,76) = 1.044$, $p = 0.310$, partial $\eta^2 = 0.014$, in the model. There was no significant main effect of sex in the model, $F(1,76) = 0.017$, $p = 0.897$, partial $\eta^2 < 0.001$. The effect of follow-up status (returner or non-returner) was not significant in the model that included the effects of age and NART score, $F(1,76) = 0.582$, $p = 0.448$, partial $\eta^2 = 0.008$. There was no significant interaction between sex and follow-up status, $F(1,76) = 0.220$, $p = 0.640$, partial $\eta^2 = 0.003$.

The estimated marginal mean anxiety score, adjusted for age and NART score, was 8.1 (SE 0.5) for the returners and for the non-returners was 7.2 (SE 1.0). The anxiety levels of the returners and non-returners in the spinal control group, as measured by the Hospital Anxiety and Depression Scale at baseline, did not differ significantly.

Hospital Anxiety and Depression Scale – Depression Score

Brain Tumour Group

The mean depression score for the 63 returners in the brain tumour group on the Hospital Anxiety and Depression Scale was 4.4 (SD 4.3) and for the non-returners was 5.0 (SD 3.9).

General linear modelling, with the depression score entered as the dependent variable, revealed no significant main effect of either of the covariates age, $F(1, 109) = 0.008$, $p = 0.927$, partial $\eta^2 < 0.001$; or NART score, $F(1,109) = 0.813$, $p = 0.369$, partial $\eta^2 = 0.007$. Sex had no significant main effect in the model, $F(1,109) = 2.443$, $p = 0.121$, partial $\eta^2 = 0.022$. Follow-up status (returner or non-returner) had no significant main effect in the model that included the effects of the covariates age and NART score, $F(1,109) = 0.287$, $p = 0.593$, partial $\eta^2 = 0.003$. There was no significant interaction between sex and follow-up status, $F(1,109) = 0.804$, $p = 0.372$, partial $\eta^2 = 0.007$.

The estimated marginal mean depression scores, adjusted for age and NART score, were 4.5 (SE 0.5) for the returners in the brain tumour group and 4.9 (SE 0.6) for the non-returners. The levels of depression, as measured by the Hospital Anxiety and Depression Scale at baseline, did not differ significantly between the returners and non-returners in the brain tumour group.

Spinal Surgery Group

The 65 returners in the spinal group scored a mean of 5.3 (SD 3.7) on the depression scale of the questionnaire and the 17 non-returners scored a mean of 5.4 (SD 3.2).

The general linear model, with depression score as the dependent variable, revealed no significant main effect of the covariate age on depression scores, $F(1,76) = 0.444$, $p = 0.507$, partial $\eta^2 = 0.006$. The covariate NART score had a significant main effect in the model, $F(1,76) = 4.712$, $p = 0.033$, partial $\eta^2 = 0.058$. Participants with lower NART scores were more likely to have lower depression scores. Sex had no significant main effect in the model, $F(1,76) = 0.299$, $p = 0.586$, partial $\eta^2 = 0.004$. The effect of follow-up status (returner or non-returner) had no significant effect in the model that included the effects of age and NART score, $F(1,76) = 0.043$, $p = 0.837$, partial $\eta^2 = 0.001$. There was no significant interaction between sex and follow-up status, $F(1,76) = 0.239$, $p = 0.626$, partial $\eta^2 = 0.003$.

The estimated marginal mean depression score, adjusted for age and NART score, for the returners in the brain tumour group was 5.4 (SE 0.4) and for the non-returners was 5.1 (SE 1.0). There was no significant difference between the depression levels of the returners and non-returners in the spinal surgery group, as measured by the Hospital Anxiety and Depression Scale at baseline.

Hospital Anxiety and Depression Scale – Total Score

Brain Tumour Group

The 63 returners in the brain tumour group scored a mean of 12.3 (SD 7.8) on the Hospital Anxiety and Depression Scale, and the 52 non-returners scored a mean of 13.2 (SD 7.2).

General linear modelling, with total score on the Hospital Anxiety and Depression Scale as the dependent variable, showed there to be no significant main effect of the covariate age, $F(1,109) = 0.015$, $p = 0.902$, partial $\eta^2 < 0.001$; or the covariate NART score, $F(1,109) = 0.424$, $p = 0.516$, partial $\eta^2 = 0.004$ in the model. Sex had no significant main effect in the model, $F(1,109) = 1.881$, $p = 0.173$, partial $\eta^2 = 0.017$. The effect of follow-up status (returner or non-returner) was not significant, $F(1,109) = 0.168$, $p = 0.683$, partial $\eta^2 = 0.002$ in the model that included the effects of the

covariates. There was no significant interaction between sex and follow-up status, $F(1,109) = 1.449$, $p = 0.231$, partial $\eta^2 = 0.013$.

The estimated marginal mean score, adjusted for age and NART score, for the returners in the brain tumour group on the Hospital Anxiety and Depression Scale, was 12.5 (SE 1.0) and for the non-returners was 13.0 (SE 1.1). There was no significant difference between the total scores of those who completed session 2 follow-up and those who did not.

Spinal Surgery Group

The 65 returners in the spinal group had a mean total score of 13.3 (SD 6.9) on the Hospital Anxiety and Depression Scale and the 17 non-returners had a mean score of 12.6 (SD 4.6).

Total score on the Hospital Anxiety and Depression Scale was entered as the dependent variable in the general linear model. There was no significant main effect of the covariates age, $F(1,76) = 0.066$, $p = 0.797$, partial $\eta^2 = 0.001$; or NART score, $F(1,76) = 3.238$, $p = 0.076$, partial $\eta^2 = 0.041$ in the model. The effect of sex was not significant in the model, $F(1,76) = 0.142$, $p = 0.707$, partial $\eta^2 = 0.002$. Follow-up status (returner or non-returner) also had no significant main effect in the model that included the effects of the covariates, $F(1,76) = 0.326$, $p = 0.570$, partial $\eta^2 = 0.004$. There was no significant interaction between sex and follow-up status, $F(1,76) = 0.301$, $p = 0.585$, partial $\eta^2 = 0.004$.

The estimated marginal mean total score, adjusted for age and NART score, on the Hospital Anxiety and Depression Scale was 13.4 (SE 0.8) for the returners in the spinal control group and 12.3 (SE 1.7) for the non-returners. There was no significant difference between the returners and non-returners in the spinal group in terms of levels of distress measured by the Hospital Anxiety and Depression Scale.

Table U.1. Overview of comparisons between the performance of the returners and non-returners in the brain tumour group and spinal surgery group on each of the baseline tests.

Test	Brain Tumour Group			Spinal Control Group		
	Returners **Estimated Marginal Mean[1] (SE)**	**Non-returners** **Estimated Marginal Mean[1] (SE)**	**p value for difference**	**Returners** **Estimated Marginal Mean[1] (SE)**	**Non-returners** **Estimated Marginal Mean[1] (SE)**	**p value for difference**
Inspection Time (All Data)	115.4 (2.2)	113.1 (2.4)	0.484	121.7 (1.5)	121.9 (3.4)	0.974
Inspection Time (Valid Data)	122.6 (1.6)	117.3 (1.8)	0.030*	124.2 (1.2)	121.6 (2.6)	0.368
Rey Auditory Verbal Learning Test	60.9 (2.1)	63.3 (2.5)	0.480	71.5 (1.5)	70.1 (3.3)	0.152
Trail Making Test Part B	90.1 (3.7)	100.5 (4.2)	0.069	81.3 (2.5)	80.9 (5.5)	0.955
Verbal Fluency	30.2 (1.4)	27.1 (1.5)	0.135	35.2 (1.4)	32.9 (3.1)	0.503

Digit Symbol Coding	59.6 (2.2)	52.0 (2.5)	0.047*	68.2 (1.6)	64.3 (3.4)	0.014*
Letter-Number Sequencing	9.2 (0.4)	9.7 (0.4)	0.726	11.1 (0.3)	11.5 (0.6)	0.596
Williams Delayed Recall Test (EFIT)	9.0 (0.6)	7.2 (0.7)	0.051	5.1 (0.4)	6.0 (1.0)	0.424
Nine Hole Peg Test – Right Hand (EFIT)	14.1 (0.5)	15.5 (0.6)	0.086	13.1 (0.3)	12.5 (0.7)	0.430
Nine Hole Peg Test – Left Hand (EFIT)	15.6 (0.4)	15.2 (0.5)	0.362	14.4 (0.4)	13.9 (0.9)	0.638
Timed Ten Metre Walk (EFIT)	6.9 (0.2)	7.1 (0.3)	0.490	7.9 (0.3)	8.0 (0.7)	0.863
Hospital Anxiety and Depression Scale – Anxiety Score	7.9 (0.6)	8.1 (0.6)	0.837	8.1 (0.5)	7.2 (1.0)	0.448

Hospital Anxiety and Depression Scale – Depression Score	4.5 (0.5)	4.9 (0.6)	0.593	5.4 (0.4)	5.1 (1.0)	0.837
Hospital Anxiety and Depression Scale – Total Score	12.5 (1.0)	13.0 (1.1)	0.683	13.4 (0.8)	12.3 (1.7)	0.570

[1] Estimated marginal mean scores, adjusted for and National Adult Reading Test score and derived from general linear models.

***** Indicates significant difference between returners and non-returners.

U.4 Discussion

Comparisons of the baseline test scores of those patients in the brain tumour group who completed session 2 ('returners') with those who did not complete session 2 ('non-returners') revealed no significant differences between the two groups in terms of the demographic variables sex, age, handedness and highest level of qualification achieved. The non-returners in the brain tumour group did however have significantly poorer scores on the National Adult Reading Test (NART) than the returners. The histological diagnoses of the two groups were not significantly different. The returners had significantly higher scores than the non-returners on inspection time, but this was only the case when valid inspection time data alone was included in the analysis. Specifically, it was the female non-returners who performed significantly worse than the female returners. When all inspection time data was included, there was no significant difference between the returners and non-returners in the brain tumour group. The returners had higher (worse) scores on the Williams Delayed Recall Test from the Edinburgh Functional Impairment Tests than the non-returners, although the difference did not quite reach the conventional level of statistical significance ($p = 0.051$). The non-returners in the brain tumour group also had significantly poorer scores than the returners on the digit symbol coding measure. Comparisons of the returner and non-returners in the brain tumour group on the other baseline tests revealed no significant overall differences between the two groups. However, the female returners significantly outperformed the female non-returners on the verbal fluency test, despite no overall difference between the returners and non-returners being found.

The returners and non-returners in the spinal surgery control group were also compared. There was no significant difference between the two groups on any of the demographic variables, including age, sex, handedness, highest level of qualification achieved and NART score. In the spinal control group, there were no significant overall differences between the returners and non-returners on any of the baseline measures. However, the female returners outperformed the female non-returners on

the Rey Auditory Verbal Learning Test (RAVLT) and also on the inspection time task when only valid data was included.

Comparing the performance of the returners and non-returners in each group is important as it gives an indication of whether differential attrition may have confounded the results and conclusions drawn regarding the effects of surgical intervention. If those patients who completed session 2 follow-up (returners) were more likely to be higher functioning patients who were less symptomatic, with better scores at baseline, they may have been less likely to have worsening of existing deficits following surgery. Therefore, the results detailed in the following chapters may under-estimate the prevalence of post-operative dysfunction as measured by the test battery. Since only 54.2% of the brain tumour group and 78.8% of the spinal control group completed session 2 testing post-operatively, it is therefore particularly important to consider potential differences between those who were followed up and those who were not. Certainly, in the brain tumour cohort, 31.6% of those who received no follow-up at all (at session 2 or session 3) were not assessed as they were unable to complete the testing sessions as a result of either cognitive impairment or physical impairment post-operatively. Therefore, those individuals in the brain tumour group who had the most severe post-operative complications were not taken into account in our analyses. That the returners in the brain tumour group performed significantly better at baseline on inspection time when only valid inspection time data was included suggests that the extent of deterioration on this measure at least may have been underestimated. Since only 4 members of the healthy control group failed to complete follow-up, comparisons of the returners and non-returners were not carried out as the very small number of non-returners in this group meant that no meaningful comparisons could be made.

In both the brain tumour and spinal control groups, there were no significant differences between the returners and non-returners in terms of sex or age. This finding contradicts other studies that have found predictors for drop out to be increased age and being male (Matthews et al., 2004). However, a study of older adults in a disability prevention trial found no significant differences between those

who did and did not return for follow-up in terms of age, sex or self-perceived health at baseline assessment (Minder et al., 2002). In general, the results of studies examining the effect of sex on attrition have been inconsistent. Our finding that the sex of the participant had no effect on whether or not the participant returned for follow-up is consistent with the findings of Mihelic and Crimmins (1997) who also found no significant effect of sex on returner status in a study detailing factors associated with loss to follow-up in a cohort of participants older than 70 years. The spinal returners and non-returners did not differ significantly in terms of their performance on the National Adult Reading Test (NART) which was administered as a measure of premorbid intelligence (see Chapter 3.5.2.2). The brain tumour returners, however, had significantly better NART scores than the non-returners, suggesting that the returners had better premorbid ability than the non-returners in this group. This likely reflects the tendency for better educated individuals and those with higher intelligence test scores to not only participate in but to be more likely to complete clinical trials. For example, van Beijsterveldt et al. (2002) report predictors of drop out to be older age and being less well educated. That the spinal surgery returners did not have significantly better NART scores than the non-returners in this group could be explained by the fact that 44% of the non-returning spinal surgery cohort were classified as 'missed' (see Figure 5.1), compared with only 13% of the non-returners in the brain tumour group. Thus, a significant proportion of the non-returners in the spinal surgery group were not approached and as such they essentially did not choose to withdraw from further participation.

The returners in the brain tumour group significantly out-performed the non-returners on inspection time testing when valid data only was included and also on the digit symbol coding measure. There was a trend towards poorer performance in the non-returner brain tumour group on the Williams Delayed Recall Test from the EFITs. However, no significant differences between the two groups on any of the other measures were found. It is would therefore appear that those brain tumour patients who completed session 2 follow-up were broadly representative of the total cohort tested at baseline and this finding is inconsistent with the findings of many studies that detail selective attrition in longitudinal studies that include cognitive

assessment. In a review of such studies of elderly participants, Chatfield et al. (2005) report that people with poorer functioning and greater cognitive impairment are consistently found to be more likely to drop out. Levin et al. (2000) also found that, in a sample of patients with Parkinson's disease, patients who performed less well on a number of neuropsychological measures were the least likely to return for follow-up. More specifically, in one of the few neuro-oncological, longitudinal studies of cognition that specifically detail attrition, Correa et al. (2007) have reported that, in a cohort of low-grade glioma patients, those patients who did not return for follow-up were significantly impaired on non-verbal recall measures at baseline by comparison with those patients who completed follow-up. These authors highlight the importance of recruiting large cohorts and employing sensitive measures of cognition in future studies in an attempt to overcome this problem.

In the brain tumour group, histological diagnosis had no significant effect on whether the participant was a returner or non-returner at session 2. This is perhaps surprising given that previous studies have found patients with more malignant tumours tend to be more symptomatic than those with less aggressive disease, for example low-grade glioma patients who are often asymptomatic (Whittle, 2004). On this basis, it could be hypothesised that high-grade glioma patients would be less likely to complete follow-up due to more severe cognitive and/or physical impairment. However, Scheibel et al. (1996) report no significant effect of histology on neuropsychological test performance of brain tumour patients in the post-operative period. Patients were classified into a 'glioblastoma' and 'non-glioblastoma' group however, and this differs from the low and high-grade glioma groups in the present and other similar studies. If the low and high-grade glioma groups did not differ broadly in terms of likelihood of experiencing severe post-operative impairment, this could explain why that those patients with more aggressive tumours were no less likely to be returners than those patients with more benign lesions in the present study. Moreover, as detailed in chapter 4, high-grade glioma patients were under-represented in the total recruited brain tumour cohort. Thus, the high-grade group who did take part were likely to be a select group who had less severe cognitive impairments than did the total cohort of newly presenting high-grade glioma patients.

It is of interest that, in both the brain tumour and the spinal surgery groups, female non-returners had significantly lower inspection time scores than female returners when only valid data was included for comparison. This was also the case for the verbal fluency measure in the brain tumour group and the RAVLT in the spinal control group. In both cases, the female returners significantly outperformed the female non-returners, although there was no significant overall difference between the returners and non-returner groups as a whole. Although several studies have reported that male participants are more likely to be non-returners than female participants, no interaction between participant sex, follow-up status and baseline performance has been reported in the literature.

Given the potential for attrition to negatively affect the representative nature of results and subsequent conclusions drawn in the study, a number of steps were taken to minimise participant drop-out in the present study. Where possible, the neurosurgeon involved with the patient's care encouraged him/her to complete post-operative follow-up prior to discharge, the researcher was flexible and available out of hours wherever possible to carry out follow-up testing and the medical staff on the relevant wards were encouraged to contact the researcher when a patient was planned for discharge. However, the majority of the non-returners either declined follow-up or were too unwell post-operatively to participate further at the time of discharge (in such cases, the patient was often transferred to another hospital for further rehabilitation). Thus, despite the detailed recruitment plan, the majority of participants who were non-returners in the brain tumour group were not recruited for reasons outwith our control. However, those patients who declined and a small proportion of those who were too unwell at the time of discharge may have been willing to participate a few days later. Therefore, had the researcher been able to visit the patient in their own home to carry out the post-operative testing session this could have reduced attrition rates to some extent. It is therefore recommended that future studies should consider conducting more flexible follow-up testing sessions, preferably in a location convenient to the patient. This methodology could not be

employed in the present study however, since the inspection time task requires a specific computer monitor that is not easily portable.

Thus, although the returners and specifically the female returners in both the brain tumour and spinal surgery groups outperformed the non-returners on a small proportion of the baseline tests, there were few significant differences between the two groups in terms of demographic variables and baseline test performance. It is therefore reasonable to assume that the observed attrition in both surgical groups is unlikely to have affected the results to a great extent. In fact, the post-operative deterioration in the brain tumour group specifically may actually be under-estimated, given that several patients who did not complete session 2 follow-up did not do so as a result of post-operative complications that resulted in physical and/or cognitive dysfunction.